Diagnostic reasoning and explanation in financial models of the firm

Feelders, Adrianus Johannes

Diagnostic Reasoning and Explanation in Financial Models of the Firm

A.J. Feelders

Diagnostic Reasoning and Explanation in Financial Models of the Firm

Proefschrift

ter verkrijging van de graad van doctor aan de Katholieke Universiteit Brabant, op gezag van de rector magnificus, prof. dr. L.F.W. de Klerk, in het openbaar te verdedigen ten overstaan van een door het college van dekanen aangewezen commissie in de aula van de Universiteit op vrijdag 12 november 1993 te 16.15 uur

door

Adrianus Johannes Feelders

geboren te Schiedam

druk: wibro dissertatiedrukkerij, helmond.

Promotor: Prof. Dr. J.P.C. Kleijnen

Co-promotor: Dr. Ir. H.A.M. Daniels

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Feelders, Adrianus Johannes

Diagnostic reasoning and explanation in financial models
of the firm / Adrianus Johannes Feelders. - [S.l.:
s.n.]. - Ill.
Thesis Katholieke Universiteit Brabant Tilburg. - With
ref.
ISBN 90-9006589-X
Subject headings: knowledge based systems.

Stellingen behorende bij het proefschrift:

Diagnostic Reasoning
and Explanation in
Financial Models of the Firm

door

Ad Feelders

1. Het is opvallend dat het "principe van zuinigheid" juist bij diagnose in het bedrijfseconomische domein niet van toepassing is.

   (dit proefschrift)

2. De methode van Kosy en Wise voor het verklaren van verschillen kan tot onjuiste resultaten leiden indien een verschil mede wordt veroorzaakt door elkaar geheel of gedeeltelijk opheffende invloeden.

   (dit proefschrift)

3. De enige reden waarom de studie van neurale netwerken tot het onderzoeksterrein van de AI wordt gerekend, en vergelijkbare statistische technieken niet, is dat neurale werken voortkomen uit een naïeve analogie met de menselijke hersenen.

4. Het onderzoek op het terrein van computerschaak heeft een slechts zeer geringe bijdrage geleverd aan de ontwikkeling van de schaaktheorie. Dit is te wijten aan het gemakkelijke succes dat te behalen viel door de combinatie van krachtige zoekalgorithmen en snelle hardware.

5. De overspannen verwachtingen die men op dit moment van het onderzoek naar neurale netwerken heeft, zijn vergelijkbaar met de overspannen verwachtingen uit de begintijd van de AI. Ze zullen dan ook tot een vergelijkbare teleurstelling leiden.

6. Degenen die house-muziek veroordelen omdat nu "iedere dwaas met een computer" muziek kan maken, hechten te veel waarde aan de beheersing van een muziekinstrument, en te weinig aan creatieve muzikale ideeën.

7. De wereld is zo moeilijk te begrijpen omdat we geen enkel referentie-object hebben om haar mee te vergelijken.

   (vgl. hoofdstuk 2 van dit proefschrift)

8. Onderzoekers die veel moeten publiceren hebben geen tijd om onderzoek te doen.

9. Het veelvuldig gebruik van "shadowboxes" in wetenschappelijke publicaties lijkt vooral bedoeld om oppervlakkige theorieën meer diepte te geven.

10. Goede bedrijfseconomen verkopen hun boeken na het behalen van de doctoraalbul.

Aan mijn ouders

For you see, ladies and gentlemen, and above all, your Imperial Majesty, with a real Nightingale one can never calculate what is coming, but in this artificial bird everything is settled.
It is this way, and no other!
One can explain it; one can open it and show how it's almost human; show where the records are, and how they play and how one thing depends on another—!

Hans Christian Andersen in
„The Emperor's Nightingale“

Contents

Introduction

A generally accepted definition of diagnosis is: "finding the best explanation for observed abnormal behaviour of a system". This definition contains two concepts that are central to diagnosis, namely *abnormal behaviour* and *explanation*.

In order to determine whether a particular system, e.g. the human body or some mechanical device, behaves abnormally, one needs to have some knowledge about how a "normal" or "correct" system would behave. This knowledge may be based on experience, on a theoretical understanding of the system or on its intended functionality.

The purpose of the diagnostic process is to find an explanation for discrepancies between observed and "normal" or "correct" system behaviour. An explanation is a hypothesis that one or more abnormal states of the system have caused the observed discrepancies. In general, diagnosis can be viewed as a hypothetico-deductive process that consists of the following steps:

- formation of a set of alternative explanatory hypotheses, based on initial findings,
- further information gathering, in order to discriminate between competing hypotheses,
- revision of the set of hypotheses to account for new data.

Whether an explanation for the abnormal behaviour of a system is adequate, has to be judged in relation to the subsequent therapy phase, since ultimately the objective will be to "cure" the system. In general, an explanation is adequate if it is detailed enough to decide which therapeutic action should be taken.

Medical diagnosis is probably the best known instance of the general diagnostic task. Relating medical diagnosis to the general description given above, the following "mapping" emerges. The system to be diagnosed is the human organism or a specific part of it. Medical knowledge about this system is incomplete, although particular subsystems may be well understood. Normal behaviour is defined as the behaviour of a "healthy human organism". The presence of a particular disease usually serves as an explanation for observed discrepancies between actual and normal behaviour. A disease can refer to either structural, functional or biochemical deficiencies in the human organism [dVR78]. The following example illustrates this. The normal human body temperature is 37.8° C. If for a particular patient, a temperature of 39° C is measured, a discrepancy between actual and normal behaviour has been discovered; this discrepancy is called "fever". A possible explanation of the fever is that the patient has some infection. There are however other plausible explanations, so additional observations are needed to find the correct explanation. If it turns out that it is indeed an infection that caused the fever, the physician may require a more detailed specification of the cause, i.e., the precise type of infection, in order to select an appropriate therapy.

Another well known instance of diagnosis is the diagnosis of technical devices. By technical devices we mean simple electronic circuits as well as more complex devices such as a car or a CD-player. In this case the system to be diagnosed, is the device considered. Human knowledge of such a device is fairly complete, since it was designed and built by man. Normal behaviour is usually derived from the intended functionality of the device. For instance, if the key of a car is turned, the car is expected to start. If the car actually fails to start upon a particular occasion, there is a discrepancy between actual an normal behaviour. One plausible explanation for this discrepancy is that the battery is not loaded, but there are other plausible explanations, so further information will be needed to discriminate between the competing explanatory hypotheses.

A third instance of diagnosis, that is actually the topic of this thesis, is the diagnosis of financial and operational performance of a business company. Diagnosis of business performance is a task that can occur in a number of different settings. It is, for example, an essential phase in the managerial decision-making process ([EN68, MRT76, Bon72]). The system to be diagnosed is a particular business company. Normal behaviour is usually defined by goals that have been set by management. In a study of Pounds

([Pou69]), it was found that managers use four types of models to define their goals: historical models (i.e. past performance), planning models (e.g. budgets), models of others (other departments, superiors), and models from the environment of the organisation (e.g. industry averages). Historical models are based on the assumption that recent past experience is the most important information for estimating the short term future. Planning models contain projections of operating variables for the coming period(s), or may contain budget values for particular cost categories.

Diagnosis is preceded by the problem identification phase, in which data are compared in order to detect discrepancies between actual performance and the goals that were defined. If a discrepancy is significant, it is viewed as a symptom that must be explained. A major issue in the comparison activity is the specification of the degree of deviation from a goal that will be allowed.

Emory and Niland ([EN68], p. 50) define managerial diagnosis as "gathering information and analyzing it to identify the limiting factors that are missing or that must be changed in order to achieve an objective". They make a distinction between diagnosis in *control* situations, where these limiting factors are usually matters of difficulty, and diagnosis in *planning* situations, where the limiting factors are more likely to mean conditions of opportunity for improvement. The limiting factors of these causes are in turn found by comparing actual performance with implicit or explicit goals. This process of identifying limiting factors should be repeated, until a level has been reached that permits corrective action. To give a simple example, suppose that a company has set a target of making a profit of fl. 1000,000.- in 1994. This target level specifies the "normal" behaviour in the diagnostic process. Suppose that the actual profit made by the company in 1994 turns out to be fl. 900,000.- Hence, a discrepancy between normal and actual behaviour is observed. This discrepancy may be explained by the fact that the total costs in 1994 were actually higher than was foreseen when the target for 1994 was set. This is probably not an adequate explanation, because it is not detailed enough to select a particular "therapy". A further specification of which particular costs were higher than planned, is needed to be able to take measures to reduce the observed discrepancy.

At this point we would like to make a distinction between two interpretations of the term *diagnosis*. The most common interpretation is that of diagnosis as finding *singular causes*. For example, it is claimed that in a particular case it was a decrease in advertising expenditures that caused sales to fall. A claim of singular causation is usually justified by a causal regularity, for example the causal regularity "a decrease in advertising *ceteris paribus* leads to a decline in sales". The regularities that are used to justify claims to singular causation, are assumed given in this interpretation. Another interpretation of diagnosis that is occasionally encountered, is that of finding *causal regularities* ([EH82, EN68]). In this case the objective is to induce causal regularities from a data set. A well known definition of causality originating from econometric data analysis is due to Granger ([Gra80]). In this thesis, however, we will exclusively be concerned with diagnosis as finding singular causes.

In Artificial Intelligence (AI) research, much attention has been given to the formalisation and automation of diagnostic reasoning. One of the earliest AI programs that was shown to reach decisions at the level of an expert in the field, was the medical expert system MYCIN, developed in the 1970s. ([BS84]). The MYCIN program provides consultative advice on *diagnosis* and *therapy* for infectious diseases, in particular bacterial infections of the blood. Since then, a host of other AI programs in the area of medical diagnosis has been developed. The earlier of these programs based their reasoning on experiential knowledge elicited from domain experts. Such knowledge consists for a large part of empirical associations, relating a pattern of symptoms, i.e. abnormal behaviour, to particular disorders. In the AI literature such a body of empirical associations is often called *shallow* knowledge, because it does not contain much information about the causal mechanisms underlying the relationship between disorders and symptoms. The most popular knowledge representation formalism for shallow knowledge is the production rules formalism. The predominant use of shallow knowledge in AI programs is associated with a number of shortcomings, such as unsatisfactory explanation capabilities and brittle problem solving ([Ste88]). These shortcomings have initiated research efforts to employ *deep* knowledge for diagnosis. The latter approaches are also indicated by the term *model based*. An important assumption in model based approaches to diagnosis is that "shallow" expert rules usually turn out to be specialized pre-compiled statements that are, in fact, derivable from the underlying theory. In model based diagnosis, however, the underlying theory is explicitly represented in the program ([AH86, CS83]). This leads to better explanation capabilities and more robust problem solving behaviour.

Model based diagnosis methods typically make use of one of the following types of models:

- a *structural model*, specifying the system's components and their interconnections,
- a *causal model* of the system, identifying the consequences of faulty system states.

Methods employing models of the first kind are said to perform diagnosis from *first principles*, because they use little or no experiential knowledge to generate hypotheses which account for data. Such models describe the *correct* behaviour of the system to be diagnosed. This approach seems to be applicable to the diagnosis of simple physical devices, where design descriptions could be used directly for diagnostic reasoning. Diagnosis based on first principles certainly eliminates some of the shortcomings associated with AI-systems that primarily contain shallow knowledge ([Jac89]):

- Given a system description, the program designer is able to short-cut the process of eliciting empirical associations from a human expert.
- Since only knowledge of *correct* system behaviour is required, the method is able to diagnose faults that have never occurred before.

However, in some real-world domains the application of diagnosis from first principles is nearly impossible ([CT90]). For example, in the medical domain, it would be far too complex to describe the structure and behaviour of a physiological system in the form required by first principles approaches. It is therefore not surprising that another approach to deep diagnostic reasoning, based upon causal models, is applied in the medical domain ([PSS84, WKAS78, dVR78]). The common characteristic of these AI-systems is that knowledge is represented by a causal model that specifies the possible states of the system to be modelled, and their cause-effect relationships.

IF the patient has fever AND
the patient has a stiff neck AND
the patient has undergone neurosurgery
THEN the best therapy is gentamycine

Table 0.1: "Shallow" rule from MYCIN

To illustrate the gradual difference between deep and shallow knowledge, we consider a production rule from MYCIN, displayed in table 0.1. A model based system would represent the underlying causal relations that justify this rule: meningitis causes an infection and a stiff neck; the infection causes the patient to have fever. If the patient has undergone neurosurgery, then the most probable cause of the infection is the bacteria "staphylococcus". The best antibiotic to treat an infection by "staphylococcus" is gentamycine. Since this underlying knowledge is not represented in the rule ("compiled away"), an explanation of its conclusion - which typically displays the rule and points at the truth of its antecedents - will be necessarily incomplete and may even be confusing to a non-expert user. Figure 0.1 gives an overview of types of knowledge representation for diagnosis within AI.

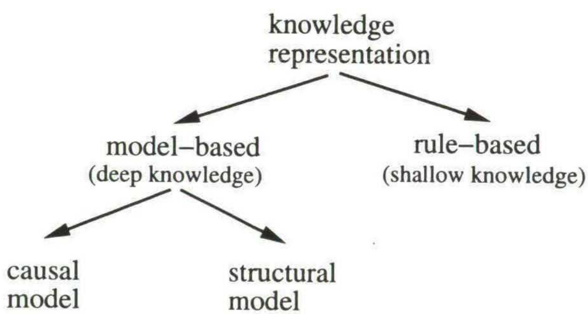

Figure 0.1: Knowledge representation for diagnosis in AI

It is important to note that AI research on diagnostic reasoning has, either explicitly or implicitly, almost exclusively been concerned with the domains of medical diagnosis and diagnosis of technical devices. The objective of this thesis, however, is to provide a formalisation of diagnostic reasoning in an area that has received only marginal attention within AI, namely the financial and operational performance of a business company. Furthermore, we will provide an implementation of this formal description in a computer program.

The merits of the research presented in this thesis are twofold. Firstly, it provides a formal analysis of the task of diagnosis of business performance. To quote Mintzberg ([MRT76], p.274):

> *Diagnosis is probably the single most important routine (in managerial decision making), since it determines in large part, however implicitly, the subsequent course of action. Yet researchers have paid almost no attention to diagnosis, preferring instead to focus on the selection routines, which often appear to be just a trimming of the overall decision process.*

Secondly, it extends the current research on model based diagnosis within AI. It will be shown that the domain of business performance requires a knowledge representation formalism and diagnostic reasoning methods, that differ from extensively investigated domains such as medical diagnosis and diagnosis of technical systems.

This thesis is organized as follows. In chapter 1 we review approaches to model based diagnosis in the medical and technical domains, and discuss three systems that are relevant to model based diagnosis in the business domain. The purpose is to compare the domains in order to determine whether reasoning and knowledge representation from the medical and technical domains can be transferred to the business domain. We argue that this is not the case. The three approaches to diagnosis and explanation of business performance discussed in chapter 1, provide a starting point for the development of a suitable knowledge representation and reasoning formalism for this task.

The notion of *explanation* is central to diagnosis. In chapter 2, we develop a new model of explanation that meets the requirements of the financial and business domain. This model serves as the basis for the formal framework for explanation generation and diagnostic reasoning that is developed in chapters 3 and 4.

In chapter 3, we describe a knowledge representation formalism for diagnosis of business performance. This formalism distinguishes between a norm model, which defines the "normal" behaviour of a company, and a business model. The business model contains both *quantitative* and *qualitative* relations among model variables. We argue that knowledge of qualitative relations among variables plays an important role in economic diagnostic reasoning.

In chapter 4 we develop new diagnostic reasoning and explanation techniques in which the norm model and business model play an important role. We distinguish between two different diagnostic situations: one involving complete information about the values of variables in the business model, and one involving incomplete information. For both situations we define the concepts of explanation and diagnosis.

In chapter 5 we discuss the implementation of the formal framework described in chapters 3 and 4. For the complete information case we use the logic programming language Prolog; for the incomplete information case we use the constraint logic programming language CHIP. In chapter 6, we summarize the results of this thesis and present our final conclusions.

In appendix A we present a case study, involving the analysis of problem-solving protocols of a financial analyst at a large Dutch bank. The results of this analysis have contributed substantially to the basic ideas underlying the formal framework for diagnosis and explanation that is developed in this thesis.

Chapter 1

Diagnostic Reasoning

1.1 Introduction

In this chapter we discuss formalizations of model based diagnosis in three different domains. The restriction to model based approaches has two reasons. Firstly, as we have already noticed in the introduction, model based approaches have advantages over AI systems based on "shallow" knowledge: the former have better explanation capabilities and more robust problem solving. Secondly, the intended application domain - diagnosis of business performance - has features that suggest that model based reasoning is a more viable approach. One reason is that, as far as financial aspects are concerned, the underlying model consists largely of simple arithmetic equations. Furthermore, causal reasoning is a common phenomenon in managerial decision making ([EN68, Bon72]).

We start with the description of two domains that are central to most AI research on diagnosis: diagnosis of technical devices and medical diagnosis (sections 1.2 and 1.3). In section 1.4 we discuss three approaches to explanation and diagnosis of business performance.

The objective is to characterize and compare diagnostic reasoning in the three different areas. The lessons drawn from this comparison will serve as a basis for knowledge representation and diagnostic reasoning in business and finance that is discussed in chapter 3 and 4.

1.2 Diagnosis of technical devices

Figure 1.1, taken from Davis and Hamscher ([DH88]), shows the basic paradigm of model based diagnosis. There is some actual device to be diagnosed, and its behaviour is observed. Furthermore there is a model of the device that can make predictions about its intended behaviour. Any difference between observed and predicted behaviour is called a discrepancy or symptom. Since it is assumed that the model describes the correct behaviour of the device, all discrepancies between observation and prediction are due to defects in the actual device.

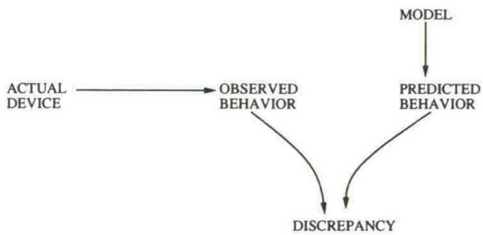

Figure 1.1: Model based diagnosis

Typically, the model used for diagnostic purposes consists of:

- a description of the internal structure of the device, i.e. its components and their interconnections, and
- a description of the behaviour of each component.

Such a model will usually be available from the design process of the device. Observations of the device are typically measurements of its inputs and outputs. The diagnostic problem is to determine which of the components could have failed in a way that explains all the discrepancies observed. To illustrate this we shall next discuss the device depicted in figure 1.2.

This device, sometimes referred to as the polybox, is a paradigmatic example in the research on model based diagnosis. It consists of the three multipliers M_1, M_2 and M_3, and the two adders A_1 and A_2. The types of inputs and their values are shown on the left-hand side of the figure; the outputs are shown on the right-hand side. The correct behaviour of each separate component can be described by stating the relation between its inputs and its output; for example, the behaviour of M_1 is given by the equation

$$\mathrm{out}(M_1) = \mathrm{in1}(M_1) \times \mathrm{in2}(M_1),$$

where $\mathrm{out}(M_1)$ refers to the output of M_1, and $\mathrm{in1}(M_1)$ and $\mathrm{in2}(M_1)$ refer to the first and second input of M_1 respectively. In the example of figure 1.2, the predicted output of A_1 ($\mathrm{out}(A_1) = 12$) deviates from its observed output ($\mathrm{out}(A_1) = 10$), whereas the output of A_2 is correct, i.e. corresponds to its predicted output.

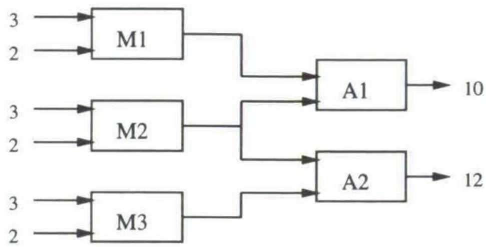

Figure 1.2: Polybox

To obtain a precise definition of the notion of diagnosis, we adopt Reiter's ([Rei87]) description, which uses first-order logic. The system to be diagnosed is a pair (SD, COMPONENTS) where

1. the *system description*, or SD, is a set of first-order sentences, and

2. the *system components*, or COMPONENTS, is a finite set $\{c_1, \ldots, c_n\}$ of constants.

Hence, for the polybox example we have COMPONENTS $= \{M_1, M_2, M_3, A_1, A_2\}$, and the following system description:

$$\mathrm{MULTIPLIER}(x) \wedge \neg\, \mathrm{AB}(x) \supset \mathrm{out}(x) = \mathrm{in1}(x) \times \mathrm{in2}(x),$$

$$\mathrm{ADDER}(x) \wedge \neg\, \mathrm{AB}(x) \supset \mathrm{out}(x) = \mathrm{in1}(x) + \mathrm{in2}(x),$$

$$\mathrm{MULTIPLIER}(M_1),\ \mathrm{MULTIPLIER}(M_2),\ \mathrm{MULTIPLIER}(M_3),$$

$$\mathrm{ADDER}(A_1),\ \mathrm{ADDER}(A_2),$$

$$\mathrm{out}(M_1) = \mathrm{in1}(A_1),\ \mathrm{out}(M_2) = \mathrm{in2}(A_1),$$

$$\mathrm{out}(M_2) = \mathrm{in1}(A_2),\ \mathrm{out}(M_3) = \mathrm{in2}(A_2).$$

The system description describes how the system components *normally* behave by appealing to the distinguished unary predicate AB($\cdot$), whose intended meaning is "abnormal". Thus, the first sentence in the system description states that a normal multiplier's output is the product of its two inputs. The remaining sentences have obvious interpretations.

The observations OBS are represented by a set of first-order sentences. For the polybox example the observations are

$$\text{in1}(M_1) = 3, \text{in2}(M_1) = 2, \text{in1}(M_2) = 3, \text{in2}(M_2) = 2, \text{in1}(M_3) = 3,$$
$$\text{in2}(M_3) = 2, \text{out}(A_1) = 10, \text{out}(A_2) = 12.$$

A diagnosis is called for, if an observation conflicts with what the system description predicts should happen if all its components were behaving correctly. $\{\neg AB(c_1), \ldots, \neg AB(c_n)\}$ represents the assumption that all system components behave correctly. Hence, such a conflicting observation can be represented by stating that

$$SD \cup \{\neg AB(c_1), \ldots, \neg AB(c_n)\} \cup OBS$$

is inconsistent. A diagnosis is a conjecture that some minimal set of components are faulty, such that consistency is restored. A set S is called minimal with respect to a property, if S has this property, but no proper subset of S does. Stated formally, a diagnosis for (SD, COMPONENTS, OBS) is a minimal set $\Delta \subseteq$ COMPONENTS such that

$$\text{SD} \cup \text{OBS} \cup \{\text{AB}(c) | c \in \Delta\} \cup \{\neg\ \text{AB}(c) | c \in \text{COMPONENTS } - \Delta\}$$

is consistent. Single-fault diagnoses are generally judged more likely to be correct than multiple-fault diagnoses, because one normally expects components to fail independently of each other. Therefore some diagnostic systems, e.g. DART ([Gen84]), make the assumption that there is only a single fault. In general, the assumption that components fail independently, leads to the preference of diagnoses with minimal cardinality. This preference criterion underlies the set covering model of Reggia et al. ([RNW84]).

In the polybox example there are four competing diagnoses: $\Delta = \{A_1\}, \{M_1\}, \{M_2, M_3\}, \{M_2, A_2\}$. Notice that if one were to make the "single fault assumption", the set of competing diagnoses would be restricted to $\{A_1\}$ and $\{M_1\}$. Since there is no unique diagnosis, it is required to make additional observations. Such an additional observation is called a measurement, or MEAS. Discrimination between competing diagnoses is based on the fact that different diagnoses may lead to different predictions of (as yet unobserved) system behaviour. A diagnosis Δ *predicts* a first-order sentence Π iff

$$\mathrm{SD} \cup \mathrm{OBS} \cup \{\mathrm{AB}(c) | c \in \Delta\} \cup \{\neg\ \mathrm{AB}(c) | c \in \mathrm{COMPONENTS} - \Delta\} \models \Pi.$$

Assuming that the components of Δ are faulty, and the remaining components are all functioning normally, system behaviour Π must hold.

In the polybox example, diagnosis $\{M_1\}$ predicts $\mathrm{out}(M_1) = 4$, and diagnosis $\{A_1\}$ predicts $\mathrm{out}(M_1) = 6$. Since these two diagnoses make conflicting predictions for $\mathrm{out}(M_1)$, it is possible to discriminate between them by measuring the output of M_1. In general, assuming that every measurement is of equal cost, the objective is to determine the actual diagnosis using a minimum number of measurements. The diagnostic process finishes as soon as there is only one diagnosis left for (SD, COMPONENTS, OBS $\cup$ MEAS).

1.3 Medical diagnosis

In the preceding section it was shown that for the domain of technical devices, there are logical theories of diagnosis. These theories assume the availability of a logical description of the structure of the system to be diagnosed, and of the normal behaviour and interaction of its components.

However, the application of this approach to diagnosis in the medical domain as well as many other real-world domains, is problematic. Describing the structure and behaviour of a physiological system in a form required by the "first principles" approach is far too complex. Consequently, model based approaches to diagnosis in the medical domain usually apply so called causal models. Causal models represent possible states of the

	node-label		
edge-sign	inc	dec	std
+	inc	dec	std
−	dec	inc	std

Table 1.1: Determination of the label of antecedent nodes

modeled system, their cause-effect relationships, and their relationships with the external world ([PSS84, dVR78]). These models can be characterised as models of behaviour without structure: they describe how one physiological event causes another. Usually such models represent the *faulty*, or in medical terms *pathophysiological*, behaviour of the system, in contrast to the models of correct behaviour that are dominant in the technical domain. As a typical example of such a causal modeling approach we give a short description of a system developed by de Vries Robbé ([dVR78]).

The model describing the disease process is a signed digraph (directed graph). Nodes in the graph represent disease characteristics and the edges represent possible causal and empirical relations between these characteristics. An edge between nodes is directed from the cause-node to the effect-node. A disease characteristic is either a variable, e.g blood pressure, or a condition, e.g. headache. With each edge a sign is associated as follows. A "+" is associated with an edge, if a change in the cause makes the effect change in the same direction. Conversely, if a change in the cause makes the effect change in the opposite direction, a "−" is associated with an edge. The symptoms that are found in a specific patient are represented by labeling the corresponding nodes in the graph. If a node in the graph represents, say blood pressure, then the symptom "increased blood pressure" is represented by the label "inc" and "decreased blood pressure" by the label "dec". If a node represents a condition, say headache, then finding this symptom is represented by labeling that node "inc". A "dec" label is not possible for a condition. A node labeled "std" represents either a variable that has remained steady, or a condition that did not occur. Edges between points represent relations that could be present in some specific case but do not always have to be.

A labeled node corresponding to a symptom that is found in a specific case, is called a search node. A search node is a node labeled "inc", "dec" or "std" in the signed digraph.

The set of search nodes is called the search set.

The possible causal chains that explain a specific symptom, are found by constructing the antecedent graph of this node. Using the label of the search node and the relevant edge-signs, the labels of the antecedent nodes are determined according to the rules given in table 1.1.

A simple example will serve to illustrate this procedure. In figure 1.3(A) a causal model, representing the pathophysiological process caused by parathyroid adenoma, is given. Suppose the findings for a specific patient are that serum bicarbonate has decreased and serum phosphate has decreased. This is represented by the search set {serum bicarbonate^{dec}, serum phosphate^{dec}}. Figures 1.3(B) and 1.3(C) represent the antecedent graphs for serum phosphate^{dec} and serum bicarbonate^{dec} respectively.

As a second step the intersection of the antecedent graphs is taken (see figure 1.3(D)). Now the "least causal" node of this intersection, i.e. parathyroid hormone level^{inc}, is called the *minimal cause* of the search set. A minimal cause replaces the original search set for the further diagnostic process. By taking the *intersection* of the antecedent graphs of the symptoms, the diagnostic algorithm implicitly gives preference to *simple* explanations, involving a minimum number of causes for the observed symptoms. We have already observed a preference for the simplest explanation in the work on technical diagnosis discussed in the previous section. This preference is based on the fact that in the domains concerned, *normal* behaviour is at the same time the *most likely* behaviour.

In case several minimal causes are found for the same set of symptoms, the symptoms can be explained by each of the minimal causes. In that case, further observations will have to point out which minimal cause is present in the situation under study, or these observations will have to show that the same symptoms are the result of several disease processes. The diagnostic procedure proposed by De Vries Robbé is more elaborate than this example might suggest. For details we refer to ([dVR78]).

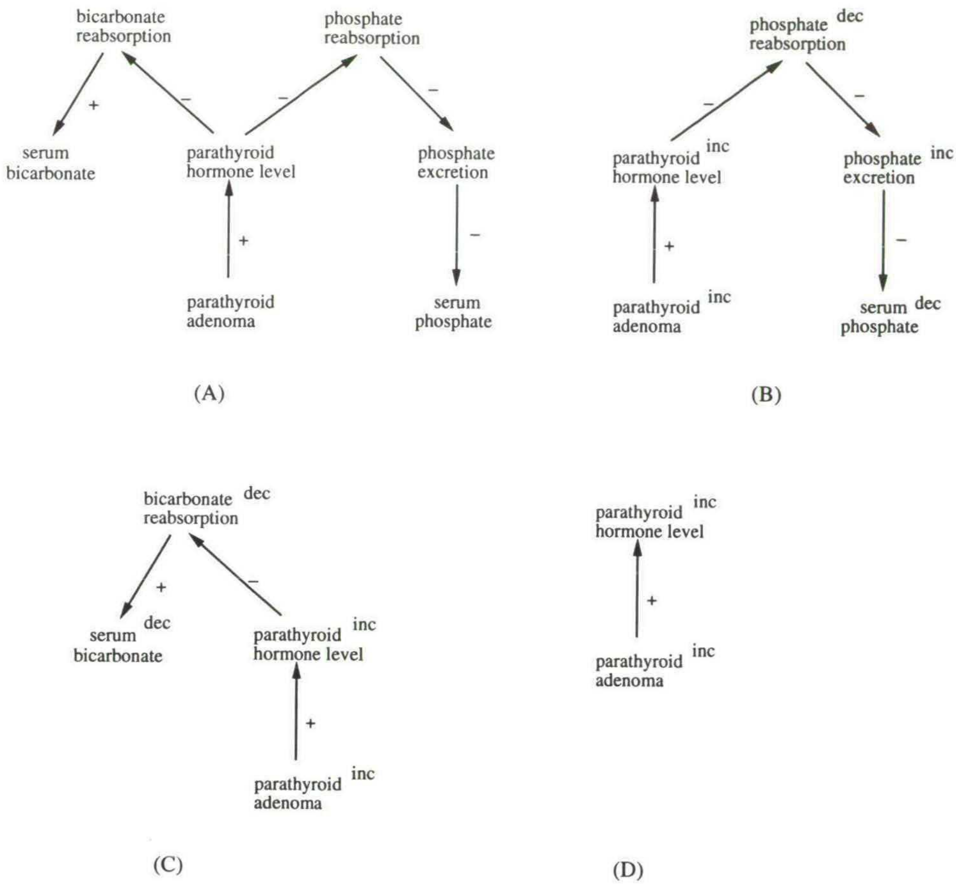

Figure 1.3: Medical diagnosis example

1.4 Diagnosis in business and finance

In this section three systems for model based diagnosis and explanation of business performance are discussed. The objective is to compare them to each other and to model based technical and medical diagnosis systems. Before we start this discussion it is necessary to state the research motivations from which the several approaches originated. None of the systems discussed in this section was presented as an AI-program for business diagnosis, but all three show a number of aspects of the problem and attempts at solutions.

The system of Kosy and Wise, described in section 1.4.1, is part of ROME (Reasoning Oriented Modelling Environment), and is used within ROME as part of a natural language query and answer system for management.

The system of Courtney et al., described in section 1.4.2, is intended as a Decision Support System (DSS) for management. In this system, interactive model construction is an important part of the diagnostic process.

Bouwman's program, described in section 1.4.3, originated from research in cognitive psychology and aims at simulating the problem solving behaviour of human financial analysts.

Finally, we should mention that the work of Rauh ([Rau88]), is also concerned with financial diagnosis of companies. However, since he takes a rule-based approach, his work lies outside the scope of this study.

1.4.1 Explanation in financial models

The work of Kosy and Wise ([KW84, Kos89]) is concerned with the explanation of model results to the users of financial models. A financial model is a representation of the activities of a business in terms of quantitative relationships among financial variables. The time span encompassed by such a model is usually divided into time periods. Kosy and Wise distinguish two kinds of explanations of model results. The first kind concerns the *value* of a model variable. In order to explain the value of a variable, it is sufficient to display the formula that was used for its computation, together with the values substituted for the variables in that formula. The second kind of explanation hinges on the comparisons the user makes between values. Values that may be compared are, for example:

- actual against historical values of a variable, and
- budgeted against actual values of a variable.

What is required then, is an explanation of the *difference* between two values. This kind of explanation is particularly relevant to diagnostic reasoning, since the purpose of diagnosis is to explain the *difference* between actual and normal behaviour of a system.

Therefore we give a description of the explanation procedure proposed by Kosy and Wise.

For ease of exposition it is assumed that a difference between period t and $t-1$ is to be explained. The results, however, also apply to the explanation of other types of differences.

The strategy to explain a difference Δy between the value of some variable y in period t and its value in period $t-1$, is to find the formula that computes the value of y, and to determine the set A of variables which appear in that formula, such that A sufficiently explains the difference Δy. Suppose $y = f(a, b, c, \ldots)$ is an equation in the financial model. Then

$$\Delta y = f(a_t, b_t, c_t, \ldots) - f(a_{t-1}, b_{t-1}, c_{t-1}, \ldots).$$

Let $S = \{a, b, c, \ldots\}$ denote the set of variables which appear in the formula for computing y. Furthermore let S^* denote the set $S - \{x \in S | \Delta x = 0\}$. Then clearly $A \subseteq S^*$.

In [KW84], A is declaratively defined as the *smallest subset* of S^* whose joint influence is sufficient to explain Δy. The influence of a set of variables $X \subseteq S^*$ is defined by the following expression:

$$\frac{\varepsilon(X, y)}{\Delta y},$$

where $\varepsilon(X, y) = y_t - f(Z)$, and Z is a vector consisting of

- the variables in X evaluated in period $t-1$, and
- the variables in $S - X$ evaluated in period t.

In words: $f(Z)$ computes the value y_t would have had, if the variables in X had not changed during the period between $t-1$ and t. A set $X \subseteq S^*$ is *sufficient* to explain Δy if

$$\theta < \frac{\varepsilon(X, y)}{\Delta y} < 1/\theta,$$

where θ is a number smaller than and close to 1.

In a later publication of Kosy ([Kos89]) a different, procedural, definition is given. To construct the set A, variables are selected from S^* in order of largest absolute effect ($|\varepsilon(\{x\}, y)|$) until their *joint effect* is sufficient to explain Δy.

	1983	1982	Difference ($=\varepsilon$)
Administration cost	635.36	614.19	21.17
Instruction	8326.56	7722.45	604.11
Pupil services	465.52	427.45	38.03
Health services	152.49	141.08	11.42
Transportation	1776.19	1545.73	230.45
Plant o & m	2289.51	2009.61	279.90
Fixed charges	2164.37	1850.85	313.52
Food services	30.19	41.26	-11.11
Student activities	217.52	208.76	8.75
Community services	31.62	27.85	3.77
Capital outlay	127.68	105.14	22.53
Debt service	1100.00	1486.35	-386.35
Outgoing transfer	685.64	655.66	29.98
Budget service	25.00	30.00	-5.00
Loan payments	125.00	115.00	10.00
Total Budget	18152.64	16963.47	1126.17

Table 1.2: School Board Budget Summary ($\times$ 1000)

In order to illustrate the explanation procedure we have adapted an example from ([Kos89]). This example will also show that the declarative and procedural definitions are not equivalent. The example in table 1.2, compares the 1983 and 1982 budget values for a Pittsburgh area school district. The difference to be explained is the increase in the total budget from 1982 to 1983. Since total budget equals the sum of all sub-budgets, the value of ε for each individual sub-budget is equal to the difference between its value in 1982 and 1983, see table 1.2, last column. Furthermore, since the equation for total budget is additive, we have $\varepsilon(X \cup Y, \text{total budget}) = \varepsilon(X, \text{total budget}) + \varepsilon(Y, \text{total budget})$.

If A is defined as the smallest subset X of S^* such that

$$0.8 < \frac{\varepsilon(X, \text{total budget})}{\Delta \text{total budget}} < 1.25,$$

where $\theta = 0.8$, the answer is $A = \{\text{instruction, fixed charges}\}$ with $\varepsilon(A, \text{total budget})/\Delta\text{total budget} = 0.81$.

In case the procedural definition is applied, i.e., variables are added to X in order of largest absolute effect ($|\varepsilon(\{x\}, \text{total budget})|$), the answer is $A = \{\text{instruction, debt service, fixed charges, plant o \& m, transportation}\}$ with

$\varepsilon(A, \text{total budget})/\Delta\text{total budget} = 0.92$. Hence, we conclude that adding variables in order of largest absolute effect does not necessarily yield the smallest set that gives a sufficient explanation. This is due to the fact that counteracting causes are also included in A. If counteracting causes are not included, it is not always possible to construct a sufficient explanation according to the above definition, since it may not be possible to find a set of contributing causes that jointly explain a fraction of Δy that lies between θ and $1/\theta$.

We conclude this section with a general remark on the explanation procedure discussed, which applies to both the declarative and the procedural interpretation. It should be noted that the explanation of the difference between two values of a variable y, is interpreted as an attempt to explain the precise quantitative difference Δy. This conclusion can be drawn from the fact that the joint influence $\varepsilon(X, y)$ of a set of variables X on the change in y, is divided by the actual difference Δy. If the ratio of the joint influence and the actual change is sufficiently close to 1, X is said to provide a sufficient explanation. Because an explanation of the precise quantitative change in y is attempted, the set A may in general contain both variables that contributed to the change in y as well as variables that counteracted the change in y. In the presentation of an explanation to the user, it is clearly indicated whether a variable had a counteracting or contributing effect. A more thorough discussion of this issue, and related issues, is given in section 4.2.1.

1.4.2 A structural modeling approach

In this section we discuss a Decision Support System (DSS) for managerial problem diagnosis described by Courtney et al. ([CPAM87, AMCP88]). In this system, weighted acyclic digraphs represent knowledge of causal relations. Weighted digraphs, sometimes called *structural models*, are directed graphs with numbers assigned to edges. Thus an edge from node x to node y, denoted by (x, y) has a number $w(x, y)$ assigned to it. In the diagnostic system of Courtney et al. the nodes in the digraph represent financial and operational variables. The weight $w(x, y)$ has the following interpretation: whenever variable x changes with u units, then (*ceteris paribus*) variable y changes with $u \times w(x, y)$

units. Hence, $w(x, y)$ represents the strength of a causal relation between two variables.

The user selects the variables in the model that have to be monitored by the system. For these variables the user specifies bounds. These bounds indicate the allowed changes for those variables from one period to the next. If the observed change exceeds one of the bounds, the variable and its consecutive values are added to a list of symptoms.

Whenever there are any symptoms, the system enters the "interactive diagnosis" phase. In this phase two kinds of analyses can be performed. The first kind is the computation of the change in a selected symptom. This change is computed using the weighted digraph model and consecutive values for the model variables. The change of variable x_i between times $t-1$ and t is computed as follows:

$$\Delta' x_i = \sum_j w(x_j, x_i) \Delta x_j$$

If the computed change ($\Delta' x_i$) is close to the observed change (Δx_i) in the value of the selected symptom, then a diagnosis may be obtained; otherwise the model does not represent the problem domain faithfully and an accurate diagnosis is not possible. This latter situation would be a problem of *model building*, which is however not our concern in this study. The analysis is only one level deep: it takes into account only those variables that directly influence the symptom.

A basic notion for the second type of analysis is that of the *causation tree* of a symptom variable. The causation tree[1] for a variable x is the subgraph generated by the antecedent set $Q(x)$ of x. This type of analysis constructs paths from terminal nodes in the causation tree to the selected symptom s. Just like a weight is assigned to an edge, one can assign a weight to a path between two nodes by multiplying the weights on the edges in the path. Suppose $C = \{c_1, \ldots, c_n\}$ are the terminal nodes in the causation tree for s. For every distinct path between a node c_i in C and the node s, its contribution to the change in s is obtained by multiplying the weight associated with the path by the observed change Δc_i in c_i. Suppose there are m such distinct paths with individual

[1]The term causation tree is somewhat unfortunate since this subgraph is not necessarily a tree

contributions pc_j. Then we define

$$\Delta'' s = \sum_{j=1}^{m} pc_j.$$

Note that $\Delta'' s$ is not necessarily equal to $\Delta' s$, since the computed change of intermediate nodes on paths from c_i to s might not be equal to the observed change. Although this possibility is not explicitly considered by the authors, it seems that in order to get a satisfactory explanation it is required that all observed changes in the causation tree should be fairly close to the corresponding computed changes. The paths are ranked on the basis of their contributions to explaining the problem. For example, if the problem variable has declined, the path with the most negative contribution is displayed first. Paths with a positive contribution that have actually offset the problem, are also displayed. This analysis is several levels deep: it takes paths into account that lead from terminal nodes to the symptom variable.

The methods described are clarified by an example taken from [AMCP88].

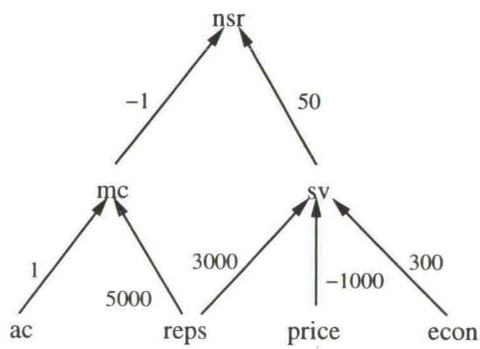

Figure 1.4: Causation tree for nsr

The symptom to be explained is a decrease in net sales revenue (nsr). In figure 1.4 the causation tree for nsr is depicted. Table 1.3 provides the relevant data for the variables in the causation tree for net sales revenue. The first analysis of the symptom yields the following result: Δ' nsr $=$ w(mc,nsr) $\times \Delta$ mc $+$ w(sv,nsr) $\times \Delta$ sv $= -1.0 \times 5000 + 50 \times -2700 = -140000$. The actual change in nsr equals -145000, and the authors consider it to be sufficiently close to call it a satisfactory explanation.

variable	period t	period $t-1$	Δ
net sales revenue (nsr)	1355000	1500000	-145000
marketing costs (mc)	95000	90000	5000
advertising costs (ac)	50000	40000	10000
number of sales representatives (# reps)	9	10	-1
price	50	50	0
economic conditions (econ)	105	104	1
sales volume (sv)	27300	30000	-2700

Table 1.3: Observed data values

PATH	WEIGHT	Δ	$\Delta\times$ WEIGHT
# reps $\rightarrow$ sv $\rightarrow$ nsr	150000	-1	-150000
adver $\rightarrow$ mc $\rightarrow$ nsr	-1	10000	-10000
price $\rightarrow$ sv $\rightarrow$ nsr	-50000	0	0
# reps $\rightarrow$ mc $\rightarrow$ nsr	-5000	-1	5000
econ $\rightarrow$ sv $\rightarrow$ nsr	15000	1	15000
Δ'' nsr			-140000

Table 1.4: Contribution of individual paths

In table 1.4 the contribution of all paths starting at terminal nodes and ending at nsr are displayed in the last column. Since nsr decreased, the path with the most negative influence is displayed first. Notice that Δ'' nsr happens to be equal to Δ' nsr, because the computed change for the intermediate variables mc and sv is equal to their actual change.

1.4.3 Bouwman's research on financial diagnosis

Bouwman ([Bou83, Bou78]) uses a qualitative model of the "typical firm" to simulate the diagnostic behaviour of a financial analyst. The financial diagnostic task required subjects to analyze various "cases". These cases were presented in the form of balance sheets, income statements, financial ratios, sales figures, and production data, for the past three years of operation of a particular firm. The subjects were asked to "make a quick evaluation of the position of the firm", and to indicate the underlying problem areas. "Thinking aloud" protocols were used to record the problem solving behaviour of the subjects. The following is a brief description of Bouwman's findings.

Market Share = (Basic Market Share − C1 × relative price) × (1.00 − C2 × lost demand) Relative price = sales price − average sales price

Table 1.5: Example of model expressions

−, ÷	down	stable	up
down	?	down	down
stable	up	stable	down
up	up	up	?

Table 1.6: Definition of subtraction and division operator

The first phase of the analysis process is *problem detection*. This is a screening activity that extracts those information items that are judged to be potentially relevant to the formulation of a diagnosis. Although the financial analysts are faced with primarily quantitative data such as balance sheets and financial ratios, they translate the series of figures into qualitative terms. The computer program developed by Bouwman uses several operators that perform *qualitative data abstraction*. Among these operators are the computation of a simple trend (increasing or decreasing), and the comparison against an industry norm. This result can get a further qualification such as "large increase" and "slightly above". After the qualitative translation, the most significant findings are selected for further processing. In general, only considerably deviating descriptions, such as "large increase" or "way below industry average" qualify as significant. The problem detection phase results in a list of significant findings. This list is the *only information* that is available for diagnostic reasoning.

The knowledge on which diagnostic reasoning is based, is represented as a set of qualitative equations that describes the functioning of a typical firm. The operators in the equations (addition, multiplication, subtraction, division, min, max) are qualitative operators, operating on the values up, down, stable, too high, and too low. Qualifications (such as "large"), that were used during problem detection are not applied during diagnostic reasoning. Table 1.5 gives an example of the expressions in the model.

Diagnostic reasoning consists of two phases: integrating significant findings, and formulating problem hypotheses. Given a particular significant finding, the program infers potential consequences through the qualitative model using the operator definitions given in table 1.6. These potential consequences are compared with the other significant findings. If a match occurs, then a causal link between the two findings is established. For example, if "sales price = up" is on the list of significant findings then, the potential consequences are generated using an equation in which sales price appears in the right-hand side. In table 1.5 this is the equation for relative price. Using table 1.6 the program infers the potential consequence "relative price = up". This potential consequence is inferred because "sales price = up" *ceteris paribus* leads to "relative price = up". In this example the *ceteris paribus* clause implies that for this inference one assumes that "average sales price = stable". If "relative price = up" also appears on the list of significant findings, then a causal link is established stating that "sales price = up" explains "relative price = up". A potential consequence of "relative price = up" is "market share = down". In this way the program determines chains and trees of related findings, called clusters, in order to focus the diagnostic process. The significant finding that explains all other significant findings in its cluster is called the *root* of the cluster.

The program uses the qualitative model of the firm to generate causes that *might* explain the root of a cluster of significant findings. It then explores the possible causes of those causes, and so on. In this way a branching tree of causes develops where each node represents a possible explanation of its higher level parent. A path from a *terminal* node to the root node of the tree represents an *explanatory path* in Bouwman's terminology. Given this collection of explanatory paths or problem hypotheses, the program ranks them on the basis of observed significant findings. Some hypotheses can be eliminated because they contradict observed significant findings. Others get a higher ranking because they are confirmed by observed significant findings. The result of this evaluation is an ordered list of explanatory paths.

Bouwman's program for financial diagnosis gives a model for some observed limitations of human diagnostic reasoning. Examples are:

- the limitation of causal chains to a maximum length,

- the restriction of the number of alternative explanations per level in the branching tree of causes.

This limitation is justified by the limited capacity of human short-term memory. These shortcomings also suggest that a diagnostic program could be of use to assist decision makers.

1.4.4 Comparison and evaluation

The goal of this section is to compare the three approaches discussed in section 1.4, and to relate them to the diagnostic modeling and reasoning approaches from artificial intelligence that we described in sections 1.2 and 1.3.

A characteristic that all three systems have in common is that they are *model based*, in the sense that a quantitative or qualitative model serves as the knowledge structure on which explanation generation and diagnostic reasoning is based. In this respect they resemble the "deep modeling" approaches in AI rather than the heuristic classification approach of, for example, MYCIN. In general the model relations express either causal or definitional relations among variables.

The quantitative approaches of Courtney et al., and Kosy and Wise are best compared to each other by looking at "one-level" deep explanations. With respect to the explanation procedure the Kosy and Wise system seems to have some advantages over the structural modelling approach of Courtney et al.:

- A more general modeling language: whereas the structural modeling approach assumes linearity, namely $\Delta' X_i = \sum_j (C_{ji} \Delta X_j)$, the system of Kosy and Wise also allows non-additive functions. Additive functions are not sufficient, even for many simple financial models.

- Kosy and Wise's explanation procedure only mentions the *significant* influences on a variable, which may prevent *information overload.*

The main common characteristic of both systems is that they require complete quantitative information in order to generate explanations. Thus diagnosis and explanation have become a matter of *information selection* rather than making "plausible assumptions". The analogous situation, in the diagnosis of some physical device, would be that it is already known beforehand which components are faulty. The program then has to point out which particular faulty components are responsible for which particular incorrect outputs.

The aim of information selection also includes the *weighing* of different influences. In Courtney et al.'s system this is done by sorting explanatory paths according to their contribution to the symptom variable. In Kosy and Wise's system the weighing of influences is effected implicitly by leaving "insignificant" influences out of the explanation, albeit that this takes place in a highly disputable manner.

Quite different is Bouwman's program. Full quantitative information is available in the cases presented to the program, but this information is not used throughout diagnostic reasoning. An artificial situation of incomplete information is created, because only *significant facts* discovered during problem detection are available to the diagnostic reasoning module. In this respect Bouwman's program has more resemblance to AI approaches where constructing explanations involves the generation of hypotheses.

Quantitative reasoning in Bouwman's system is limited to the determination of significant findings. For the generation of explanatory hypotheses, one switches to a qualitative representation, because an attempt to give quantitative explanations for observed significant findings would lead to an infinite number of hypotheses.

1.5 Conclusions

We have reviewed approaches to model based diagnosis and explanation in the technical, medical, and financial domains. Although this overview is in no sense complete, we argue that the approaches discussed are typical for their respective domains.

From a modeling viewpoint there are some notable differences. In the diagnosis of technical systems the models describe the correct or normal behaviour of the system, by specifying the normal behaviour of components and the way they are interconnected. This is a practical choice, because such models will usually be available as a product of the design process of the artifact. In model based medical diagnosis the model usually describes causal relations between incorrect or abnormal states of the system. The models used by the three financial diagnosis systems can describe both correct and incorrect behaviour.

From a reasoning viewpoint the most notable difference is that a situation of incomplete information is presupposed in medical and technical diagnosis, whereas this is certainly not the case in financial diagnosis. We have seen that explanation and diagnosis in the systems of Kosy and Wise, and Courtney et al. respectively were not concerned with forming plausible hypotheses, but with the weighing and selection of quantitative influences. The Bouwman system is an "intermediate" form. Complete quantitative information is available but is used only during the "problem detection" phase. Diagnostic reasoning takes place in a situation of incomplete information, which forces the use of qualitative reasoning in order to get a finite hypothesis space.

In the next chapter we shall develop a new model of explanation that we consider appropriate for diagnostic reasoning in the domain of business and finance. We will compare each of the three approaches discussed in this chapter to the new model and show their similarities and differences. This model will serve as the basis for the new diagnostic reasoning methods developed in this thesis.

Chapter 2

Models of explanation

2.1 Introduction

We have defined diagnosis as the problem of finding the best explanation for observed abnormal behaviour of a system. Chapter 1 reviewed a number of approaches to model based diagnosis in the medical, technical, and business domain. One of the conclusions of this review was that the economics and business domain has substantially different characteristics than other diagnostic domains, and requires the development of new knowledge representation and reasoning methods.

We reviewed three model based systems for explanation and diagnosis of business performance that are potentially relevant to the development of such new methods. We believe, however, that in order to come to a useful concept of diagnosis in the business domain, we should first make a thorough analysis of the concept that is central to diagnosis, namely *explanation*. Once this concept is clearly defined, it is only a small step to an appropriate definition of the concept of *diagnosis*.

In this chapter we provide this analysis by investigating models of explanation. In section 2.2 we discuss the well-known Deductive Nomological model of explanation. On the basis of characteristics of the business and economics domain, this model is rejected. We select a concept of explanation, so called *aleatory* explanations, that is in accordance with both the partly quantitative nature of the domain and the prevalence of causal reasoning. Aleatory explanations will serve as the basis for a formal framework for explanation and diagnosis that will be described in chapter 4. Section 2.4 compares the three approaches to explanation in business and finance, discussed in chapter 1, to aleatory explanations. It will be shown that all three approaches are related to aleatory explanations, to a certain extent.

2.2 The Deductive-Nomological model

In his "Aspects of Scientific Explanation", Hempel ([Hem65]) notes that a scientific explanation may be regarded as an answer to a why-question. Such a why question has the general form "Why is it the case that E ?" where E denotes an empirical statement specifying the explanandum. Questions of this type are called explanation-seeking why questions.

An explanation according to the Deductive-Nomological (D-N) model may be conceived as a deductive argument of the form:

$$\begin{array}{c} C_1, C_2, \ldots, C_n \\ \underline{L_1, L_2, \ldots, L_m} \\ E \end{array}$$

Here $C_1, \ldots, C_n$ are sentences describing the particular facts invoked, and $L_1, \ldots, L_m$ are the general laws on which the explanation rests. Jointly these sentences are said to form the explanans. The conclusion E of the argument is a sentence describing the explanandum phenomenon. Since the explanans is required to contain at least one lawlike sentence, the D-N model is often referred to as the *covering law* model. The lawlike sentences involved can have many different logical forms, ranging from simple universal conditional form $\forall x(F(x) \rightarrow G(x))$ to more or less complex mathematical relationships among different quantitative variables.

Consider the following economics example. In economics, theories are often expressed by mathematical relationships among quantitative variables. The equation

$$D_t = a - bP_t + cA_t$$

for example, describes a relation between the demand (D) for a particular product in period t, the price (P) of that product and the advertising expenses (A) related to it. If we want to explain the event of demand taking a particular value in period k, an explanation according to the D-N model would run as follows:

$$\frac{P_k = d, A_k = e, \quad D_t = a - bP_t + cA_t}{D_k = a - bd + ce}$$

where the equation for demand serves as the covering law and price and advertising expenses in period k as the initial conditions. The explanation shows that the value of D in period k can be *deduced* given the initial conditions and the appropriate covering law. An explanation according to the D-N model, can be represented in the format "E because X", where X contains both the initial condition(s) and covering law(s).

One of the main objections against the D-N model is its inability to account for observed asymmetries of explanation ([vF80, Hum89]). To cite a classical example ([Hum89], p. 118):

> *A flagpole of height h casts a shadow of length l. With knowledge of the length of the shadow, of the angle of elevation of the sun, and of elementary laws of geometrical optics, such as the almost rectilinear propagation of light, we can* **deduce** *the value of h. But citing the value of l does not* **explain** *the value of h. Still the deduction is a good D-N explanation.*

Similarly, in the economics example we have the valid D-N explanation:

$$\frac{D_k = d, P_k = e, \quad D_t = a - bP_t + cA_t}{A_k = \frac{d-a+be}{c}}$$

yet we would hardly be willing to accept an explanation of the level of advertising in period k by referring to the level of demand in period k and the price level in period

k. In this particular case the reason is that the advertising level in period k *causes* the demand level in period k and not vice versa. The asymmetry of explanation derives from the asymmetry of the causal relation.

Since equality is a symmetric relation from a mathematical viewpoint, it must be concluded that merely stating the lawlike sentence

$$D_t = a - bP_t + cA_t$$

does not yield an adequate description for the purpose of generating explanations. It is necessary to specify the direction in which explanations are allowed to proceed. In [Bou84] it is noted that in economic modeling every equation has a specific "reading" associated with it. Generally, in economics, the one variable appearing on the left-hand side of the equation is taken to be explained by the variables that appear on the right-hand side.

2.3 Causal Explanation

According to a causal model of explanation, phenomena are explained by giving their causes. A causal model of explanation avoids the objection against the deductive-nomological model presented in the preceding section, because causes explain effects and not the other way around.

Our exposition on causal explanation is largely based on Humphreys' notion of *aleatory* explanations ([Hum89]). Aleatory explanations have been introduced to account for probabilistic phenomena, but they are also applicable in deterministic contexts. Humphreys identifies a number of essential properties of causal explanations. Firstly, there is the *multiplicity* and *separateness* of causal influences on a given phenomenon. For example, the demand for a particular product is influenced by its price, quality, and the advertising expenses devoted to the product. The first one usually has a negative influence, whereas the second and third factors have positive influences. Usually, all of these factors will be causally effective at the same time in a particular situation. These multiple causes are separate to the extent that one is able to isolate the effect of a single factor.

The second property of causal influences is that they come in two kinds: *contributing* and *counteracting*. Within a probabilistic framework contributing causes are those causes that increase the probability of the effect and counteracting causes are those causes that lower the probability of the effect. Within a deterministic framework the distinction between contributing and counteracting causes depends on whether one requires an explanation of the effect in a *qualitative* sense or in a *quantitative* sense. The following illustration of this point is due to Humphreys ([Hum89], p.105):

> *Consider a room which is both heated and air-conditioned, and suppose that the temperature rises by* 5° *C (compared with the situation where neither is operating). Alone, the heater would raise the temperature by* 10° *C; alone, the air conditioning would lower it by* 5° *C.*

Now consider the following two claims:

1. The increase of 5° C in the temperature of the room occurred because of the input of 10,000 Btu (British thermal unit) from the heater *despite* the extraction of 5000 Btu by the air conditioning.

2. The increase of 5° C in the temperature of the room occurred because of the input of 10,000 Btu from the heater *and* the extraction of 5000 Btu by the air conditioning.

There are, in fact, two aspects of the explanandum event that need to be explained: the fact that temperature increased and the exact value of that increase. If an explanation of the increase is called for, then (1) is the correct explanation. If an explanation of the exact value of that increase is required, then (2) is the correct explanation. In the first case one explains why the temperature increased rather than decreased, or remained steady. This involves a *qualitative* description of the explanandum event. In the second case one explains why the temperature rose with exactly 5° C rather than change with some other quantitative amount. This involves a *quantitative* description of the explanandum event.

	1990	1991
D	78	84
P	8	9
A	20	40

Table 2.1: Data for demand equation

In view of the above requirements Humphreys proposes the following canonical form for causal explanations, which deviates from the "E because X" format of D-N explanations:

Event E occurred because of Φ, despite Ψ,

where Φ is a non-empty set of *contributing* causes and Ψ a (possibly empty) set of *counteracting* causes. The explanation itself consists of the causes to which Φ jointly refers. Ψ is not a part of the explanation of E proper. The role Ψ plays is to give us a clearer notion of how the members of Φ actually brought about E. Thus Ψ may be empty, in which case we have an explanation involving only contributing causes to E 's occurrence. However, if Φ is empty, then we have no explanation of E's occurrence at all. Note that in the case of quantitative deterministic explanations, such as (2) above, the "despite" clause is always empty because of the absence of counteracting causes.

Let us return to the example of the previous section, where demand (D_t) for a particular product was stated to be caused by its price (P_t), and the advertising expenses (A_t). Suppose this causal relation is quantified as follows:

$$D_t = 100 - 4P_t + 0.5A_t.$$

Suppose furthermore that the data in table 2.1 for two consecutive time periods are available. To explain why demand D increased from 1990 to 1991, the following explanation seems to be correct:

"Demand increased because advertising expenses increased,
even though price increased."

Thus $\Phi = \{\text{advertising expenses increased}\}$ and $\Psi = \{\text{price increased}\}$. In case an explanation for the exact increase of demand is required, then $\Phi = \{$advertising expenses increased, price increased$\}$ and $\Psi = \{\}$.

We did not yet thoroughly specify what counts as an event in the context of causal explanations. From the examples given thus far one can identify two kinds of events:

- variable Y has value y at time t
- variable Y *changes* value from time t to t', with $t < t'$.

An explanation of the first kind, i.e. why $Y = y$ holds, has an empty "despite" clause, since all causally relevant factors will *contribute* to Y having that exact value. For example, to explain why demand is equal to 78 in 1990, the value of both price and advertising expenses should be cited as contributing causes. For events of the second type we made a distinction between explanation of the *direction* of change (qualitative explanandum) and explanation of the *magnitude* of the change (quantitative explanandum). We will generalize this second type of event, in order to be able to explain a broader class of phenomena. The event of a variable Y changing from time t to t' can be shown to be a special case of the more general "event" of there being a difference between two values of Y evaluated with respect to different "objects". In order to make this clear, we discuss a theory of explaining differences as developed by Hesslow ([Hes84]), and show that explaining changes of Y over time is a special case of explaining the difference between two values of Y, evaluated with respect to different objects.

According to Hesslow all explanations of individual facts of the form $F(a)$ - object a has property F - involve a, sometimes implicit, comparison with other objects which lack the property in question. The following example is given by Hesslow ([Hes84], p. 87):

> *Suppose, for instance, that a fire broke out in a barn because some careless person dropped a burning cigarette in the hay. The cigarette, we may assume, was clearly a cause according to common sense but also according to all reasonable definitions of 'cause'. It was both necessary and sufficient in the circumstances; it was a part of a universally sufficient condition; it raised the probability of the fire, etc. The same, however, is also true of several other conditions such as the presence of oxygen and inflammable material*

and the absence of dampness and an automatic fire extinguishing system. All of these factors are causes in a wide sense, but none of them can explain the fire. It would clearly be absurd to try to explain the fire by pointing out the abundance of oxygen in the air.

Why then does the dropping of a cigarette in the hay seem a reasonable explanation, whereas the presence of oxygen in the air does not? This is so because we are implicitly comparing the barn that caught fire with other barns that did not burn. Since oxygen is present in all barns, this factor cannot explain the difference between those barns that burn and those that do not.

The objects of comparison are said to belong to a reference class R. The only restriction put on R is that its members must not possess the explanandum property. Hesslow's theory leads to a more detailed specification of the event to be explained. Instead of representing the explanandum as $F(a)$, it is specified further by explicitly including the reference class R. Consequently, the explanandum is a three-place relation between an object a (e.g. the barn), a property F (being on fire) and a reference class R (other barns).

$$< a, F, R >$$

The task is not to explain why a has property F, but rather to explain why a has property F *when the members of R do not.* There are several typical ways of forming reference classes, for example:

a) R as the statistically normal case.

If a certain causal condition were normal, it would occur among the members of R. In that case it could never explain the difference between the explanandum object and those in R. It follows that when R is chosen as the statistically normal, the explanatory cause must be abnormal.

b) R as the temporally normal case.

If the question is "why did the barn catch fire at this particular time", we are asking about a temporal difference, and the proper object of comparison will not be another barn, but this barn at an earlier time.

c) R as a theoretical ideal.

In many sciences it is a common procedure to use as an object of comparison a *hypothetical* object or state of affairs which is defined by some theory. Such *theoretical ideals* have the obvious advantage of providing the scientist with a constant object of comparison, thus facilitating systematization of the field covered by the theory. A typical example of such a theoretical ideal in medicine is the physiology of the healthy human organism.

Viewing explanation as explaining a difference between an object a and a reference class R, makes clear that explaining a change in a variable Y between times t and t' is just a special case of explaining a difference between two values of Y evaluated with respect to different objects, corresponding to the special case where R happens to be the "temporally normal" case. We will usually be concerned with cases in which R has only *one* member, in which case the explanandum is denoted by $< a, F, r >$.

If E is replaced by this more detailed explanandum, the following new canonical form for explanations is obtained:

$< a, F, r >$ because Φ, despite Ψ.

The following example shows that choosing different objects of comparison leads to different contributing and counteracting causes. Suppose we are studying the results of a specific business firm, called the ABC company. We are interested in this company's return on total assets:

$$\text{return on total assets} = \text{gross margin} \times \text{total assets turnover}.$$

Note that the relation between "return on total assets", "gross margin", and "total assets turnover" is not causal in the usual sense of the word. In other words, "gross margin" and "total assets" are not causally prior to "return on total assets". We do maintain, however, that "gross margin" and "total assets turnover" are *explanatorily prior* to "return on total assets". This is indicated by the fact that "total assets turnover" is the variable that appears on the left-hand side of this equation. Suppose the data in table 2.2 are available to us. We would like to explain why ABC's return on total assets is low

	ABC (1990)	ABC (1991)	industry average (1991)
return on total assets	16	12.75	16
gross margin	12	8.5	8
total assets turnover	1.33	1.5	2

Table 2.2: Data for ABC - example

in 1991. It is convenient to introduce some notation at this point. $Y^{\uparrow}$ means that the value of Y for the explanandum object is above the value of Y for the reference object and similarly for $Y^{\downarrow}$. We answer this question for two different reference objects:

a) <ABC(1991), return on total assets$^{\downarrow}$, ABC (1990)> because {gross margin$^{\downarrow}$}, despite {total assets turnover $^{\uparrow}$}.

b) <ABC(1991), return on total assets$^{\downarrow}$, industry average (1991)> because {total assets turnover$^{\downarrow}$}, despite {gross margin$^{\uparrow}$}.

We can see from this example that different contributing and counteracting causes are selected depending on the object of reference that is chosen. This phenomenon is expressed clearly by White ([Whi65], p.8):

> *If we think of the behavior of a country as abnormal when compared with its own behavior just before the puzzling event, we may explain the puzzling event in one way; but if we think of that country as behaving abnormally by comparison with other countries, we may seize upon something else as the cause of its abnormal behavior.*

Many writers have regarded the selection of causes that arises from the choice of object of comparison as a pragmatic element of explanations, thus suggesting that such a selection is non-logical and subjective. By giving a more detailed specification of the explanandum however, i.e. by explicitly stating the reference object(s), the problem of contradictory selections does not arise.

Relations between variables in the area of economics are often expressed as functional dependencies among quantitative variables:

$$Y = f(X_1, \ldots, X_m),$$

where Y is taken to be caused by (depend on) $X_1, \ldots, X_n$. We concern ourselves with the causal explanation of individual events, where the explanation rests on such a functional dependency. Depending on the level of specificity that one requires of the relation between cause and effect, different constraints should be imposed on the function f.

One may require that the cause in question contribute a precisely specified quantitative amount to a quantitative effect variable. This requirement leads to the restriction of f to additive functions. The reason is stated by Humphreys ([Hum89], p. 28) as follows:

> *So suppose that we take a quantitative approach to causation and claim that a unit change in a variable X_i caused a change of y units in Y when all other factors X_j were held constant. Then that change of y units must occur irrespective of the particular level at which the other X_j happen to be. For if not, it was not the change in X_i that caused the change of y units in Y but the change in X_i together with the prevailing level of the X_j 's.*

The assumption that the contribution of an individual variable X_i must be the same at whatever level the other variables are held constant leads to the restriction of f to additive functions:

$$f(X_1, \ldots, X_m) = \sum_{i=1}^{m} h_i(X_i) + h_0,$$

where h_0 is a constant. This follows from the fact that $\partial f / \partial X_i = g(X_i)$, i.e. $\partial f / \partial X_i$ is independent of X_j for $j \neq i$. If one requires merely that the causal factor contributes to the effect rather than counteracts it, then f does not have to be restricted to additive functions. Take for example the well known Cobb-Douglas production function in the economics literature:

$$P = \gamma L^{\alpha} C^{1-\alpha},$$

where $\gamma > 0$, $0 < \alpha < 1$, $C, L > 0$, and C denotes capital input, L denotes labour input and P denotes production volume. This production function is clearly non-additive. Still it seems legitimate to allow changes in capital as well as labour to cause a change in production volume. An increase in the amount of labour (L) results in an increase in production volume (P), irrespective of what value the other causal variable (C) has, as long as $C > 0$. The precise *quantitative value* of the increase caused by L, clearly depends on the value of C.

This is the reason why one can allow non-additive functions in qualitative claims but not in specific quantitative claims of causal contributions, unless one has effectively turned a non-additive relation into an additive relation by holding constant all but one variable, thereby implicitely limiting the causal claim to a specific set of contextual factors.

Suppose we have the following production function:

$$P = 5 \cdot C^{2/5} L^{3/5}.$$

It is possible to determine the quantitative contribution of a change in L by holding C constant at a particular level, say $C = 10$. This yields the degenerate additive function:

$$P = 5 \cdot 10^{2/5} L^{3/5}.$$

Clearly, in this situation, the quantitative causal claim that a one unit change from 8 to 9 in L leads to a change in P with 3.26 units, is warranted.

2.4 Comparison of different approaches

In this section, the three approaches to explanation and diagnosis of business performance discussed in chapter 1, are compared with aleatory explanations discussed in this chapter. This will enable us to get a "unifying" view. We will evaluate to what extent these different approaches provide explanations that coincide with the canonical format:

$< a, F, r >$ because Φ, despite Ψ.

The approach of Kosy and Wise (section 1.4.1) coincides with this canonical format, where the specification of the explanandum event $< a, F, r >$ is concerned. Let us return to the School Board Budget example in chapter 1, where the increase in total budget had to be explained. We obtain the following "instantiation" of $< a, F, r >$:

$<$ School Board (1983), Total Budget$^{\uparrow}$, School Board (1982)$>$.

In this case the choice of reference object implies that a *change* in Total Budget has to be explained. We have seen, however, that in Kosy and Wise's approach, any difference between two values of a variable can be explained, as long as these values have been generated by the same equation. A flaw in their approach, however, is that they do not make a strict *a priori* separation between contributing and counteracting causes. The set A may contain *both* contributing and counteracting causes. This "mix-up" leads to a number of difficulties that will be discussed more thoroughly in section 4.2.1. Most important is that it may cause the system to leave out significant contributing or counteracting causes from the explanation. With respect to the functional relations that are allowed to sustain explanations, Kosy and Wise allow both additive and non-additive functions. Consequently it is not always possible to make quantitative causal claims, although the measure ε is used as a "surrogate" to measure the relative strength of causes in case of non-additive functions.

In the approach of Courtney et al., r is restricted to the "temporally normal" case, i.e. the system only explains *changes* from one period to the next. Since the functional relations that are allowed to sustain explanations are restricted to *linear* functions, it is possible to make quantitative causal claims, such as: "the change Δx in x caused a change of $w(x,y) \cdot \Delta x$ in y". A clear distinction is made between contributing and counteracting influences. This distinction is also made with respect to the influence of paths.

In Bouwman's approach, r may be the "temporally normal" case or may represent industry average. Explanations contain only qualitative causal claims, such as "relative price = up" explains "market share = down". Furthermore, explanations mention only *contributing* causes. This is understandable since Bouwman's program works with incomplete information. If hypotheses concerning counteracting causes have to be made, an unmanageable number of hypotheses may arise. Providing an explanation that also mentions counteracting causes, is only appropriate if it is actually *known* that these counteracting causes occurred.

The functions that sustain claims concerning contributing causes may either be additive or non-additive. Since no quantitative causal claims are made, the use of non-additive functions does not yield any difficulties.

2.5 Conclusions

The deductive-nomological model fails to give an adequate account of explanation in the domain of business economics. It is unable to capture observed asymmetries of explanation that are due to the asymmetry of the causal relation. Therefore the causal direction of relations among variables should be represented explicitly or implicitly, in order to be able to provide adequate explanations.

Aleatory explanations have a number of characteristics that make it suitable for the domain of business economics. Firstly, explanations are allowed to proceed from cause to effect, and not vice versa. Secondly, an *aleatory* explanation of a particular event distinguishes between contributing causes and counteracting causes. This distinction is essential for the explanation of phenomena that involve quantitative variables, as is often the case in business economics. Explanations of individual events are sustained by general laws, expressed as functional relations among variables. The form of this function determines which causal claims can be made for separate variables. If the functional relation is additive, it is possible to state the quantitative contribution for each variable. If the functional relation is non-additive, such a quantitative claim cannot be made unless all but one variable are held constant. In that case the claim holds only for a given set of contextual factors.

All three approaches discussed in chapter 1 can somehow be related to aleatory explanations, but all fall short at some point. In chapter 4 we shall use aleatory explanations as the basis for explanation generation and diagnostic reasoning. Before the reasoning component is discussed, however, it is necessary to investigate what types of relations among variables are used in the business economics domain to sustain explanations. This will be done in the next chapter.

Chapter 3

Knowledge representation for diagnosis of business performance

3.1 Introduction

In chapter 2 we have presented an analysis of the notion of explanation. This has resulted in the development of a new model of explanation. The new model combines Humphreys' distinction between contributing and counteracting causes with a detailed specification of the explanandum as proposed by Hesslow. We proposed the following canonical format for explanations: "$< a, F, R >$ because Φ, despite Ψ", or when the reference class contains only one object: "$< a, F, r >$ because Φ, despite Ψ". For an example of this type of explanation, the reader is referred to section 2.3.

In this chapter, we direct our attention to the specification and representation of the knowledge that is required for diagnosis of business performance. Diagnosis of business performance is defined here as explaining the difference between the actual performance of a company, and its norm performance. In our canonical explanation format, a refers to the object that displays actual behaviour, and r refers to the object that displays norm behaviour. We identify two principal knowledge structures for diagnosis of business performance:

- Knowledge of normal behaviour: the norm model.

- Knowledge of general laws, relating variables pertaining to business performance: the business model.

The norm model specifies which reference object(s) should be used to compare a company's actual performance with. It also specifies with respect to which properties, i.e. variables, the comparison should be made. The reader may wonder why the norm model is worth discussing at all, since it is hardly addressed in most expositions on diagnosis. The reason is that the nature of norms in the business domain is considerably different from norms in diagnosis of physical devices and medical diagnosis. In these latter domains, which have largely motivated the research into diagnostic reasoning in AI, normal behaviour of the system is fairly absolute and constant through time. For simple physical devices, as we saw in section 1.2, normal behaviour can be deduced from the design description of a device.

Another way of stating the difference between the business domain on the one hand and the medical and physical devices domain on the other hand is that their objects of comparison -the healthy human organism and the intended functioning of the device- do not change over time, whereas in the business domain the proper object of comparison constantly changes. A company's performance may be considered satisfactory in one year, whereas the same performance is considered mediocre for the next year, simply because the object of comparison has changed. This is due to the fact that business organizations operate in a dynamic environment, where for example macro-economic developments to a large extent determine performance.

How does one determine which reference object(s) are appropriate for diagnosis of business performance? In the literature, one often encounters the phrase that one should not "compare the incomparable". The question that should be addressed then is: which properties are allowed to vary between objects of comparison and which are not? One could say that the choice of reference object r determines which properties are the same for a and r. Since properties that are identical for objects can never explain a difference between them, the choice of r determines for a large part the properties that are allowed to explain a difference. Thus, if one allows a jewelry store j to be compared only to other jewelry stores, e.g. $r =$ "the average jewelry store", then the nature of the product sold can never explain the difference in, for example, the gross margin between j and its object of comparison. If, on the other hand, it would be allowed to compare j to arbitrary retail stores, e.g. r = "the average retail store", then one could probably explain the difference in gross margin in terms of the relatively high markup on jewels, which is a consequence of the low turnover of goods in this branch. The question of which properties are allowed to vary between objects of comparison, is equivalent to the question of which properties are allowed to explain a difference between them.

There is one general trait of properties that are not allowed to explain a difference between the performance of business firms, namely those properties that do not reflect the performance of the *individual* firm. Macro-economic developments, for example, cannot be significantly influenced by the individual company and therefore do not reflect any performance aspect of a single company. As a consequence, temporal comparisons of a company's results are not considered appropriate if the macro-economic climate has drastically changed over the time periods that are compared. Analogously, if two companies use different asset valuation methods, their results are deemed incomparable because the valuation method used does not express a measure of performance.

The second knowledge structure that is indispensable for diagnosis is the business model. As we have already stated in the conclusions of chapter 2, explanations of specific events should always be supported by a general law. In chapter 2 these general laws were represented by quantitative functional relations. In reality, however, there may only be knowledge present of qualitatively specified laws, especially where causal relations are concerned. Therefore the business model may contain qualitative as well as quantitative relations.

In the following two sections, we discuss the norm model and business model in more detail.

3.2 The norm model

Although in all cases the object to be diagnosed is the performance of a particular company, the *viewpoint* of the diagnostician is of considerable influence on the variables

	1980	1981	1982	1983	1984
International companies	6	5	5	6	6
Trade	2	1	1	2	2
Industry	2	2	2	3	3
Miscellaneous	1	1	1	2	1
Shipping and aviation	2	2	2	1	2
All companies	4	4	4	5	5

Table 3.1: Interest coverage for Dutch companies

that are selected as measures of performance. Helfert ([Hel78]) discusses the viewpoints of management, the owners (shareholders), and lenders. These three interest groups look at different aspects of performance. The management of the firm, for example, is mainly concerned with the efficiency and profitability of operations and the effective use of capital, whereas a lender is more concerned with the firm's ability to pay back the principal and interest.

In the following, the most common "reference objects" to diagnose business performance are described.

Theoretical norm values

Through theoretical norm values one tries to establish a norm for a particular financial or operating variable that is applicable to all companies. The problem with such "universal" rules is that the variables concerned often differ significantly depending on the industry, season, policies of the firm, etc. Take for example *interest coverage*, which indicates to what extent a company is able to pay its interest expenses out of its gross profit. In the literature one often finds a norm value of 4, or even more. Table 3.1 shows the actual values of interest coverage for Dutch companies, subdivided by industry. Since, for example, all trade companies have an interest coverage that is substantially below the theoretical value of 4, this severely limits its applicability as a norm for acceptable performance for this kind of firm.

Historical norm values

In this case the norm value for a particular variable is its value in one or more previous time periods. The number of historical periods considered in the analysis should not be too large, in view of the possibility of "structural changes", such as a change in the macro-economic climate. As we have already remarked such a structural change would lead to "comparing the incomparable".

In historical comparisons the only judgment that can be made is *better* or *worse* than in the previous period. It does not enable one to say that some property is *good* or *bad* in an absolute sense. It might be the case that a company has a declining return on total assets, but that the industry on average is doing even worse.

Industry average as norm values

The industry average of companies operating within the same industry is often used as a norm for the individual company. Such industry averages can only be usefully compared for *ratios*, not for *nominal* variables. For example, knowing that the current assets for a particular firm are lower than industry average does not say much about its liquidity position unless we also know the level of its current liabilities.

The problem with any comparison between different firms is that factors such as accounting methods or the size of the firm may have a considerable influence on the results. Since these factors might not be constant among firms in the same industry, this diminishes the comparability of the firms.

Another drawback is that firms are compared to "mediocrity", i.e. industry average, and not for example to the best in their line of business. For companies that are currently below industry average, this norm may be a good objective to aim for. Firms that are above average, will probably set different goals.

Despite these possible objections, the industry average is often viewed as the *normal case* when comparing several firms in one line of business; significant deviations from this norm are viewed as a signal to look for underlying causes.

Pairwise comparison

In view of the limitations of comparing a company to the industry average it would be nice to compare one firm with an almost identical "twin". In practice it may be very hard to determine whether one company is really comparable to the other. Furthermore, if this other company is doing bad, it is not a useful object of comparison.

Plans and budgets as norm values

When the diagnostician is a firm's manager, the norm values may be the result of an *explicit* planning process. A plan may for example indicate the production to be achieved or it may contain budget values for particular expense items. A significant difference between actual and planned performance will attract the attention of management, and will lead to the search for underlying causes.

Apart from the problem of determining which norm is appropriate in a particular situation, there is the task of distinguishing between *significant* and *insignificant* differences between actual and normal behaviour. Little research has been performed on how human analysts make this distinction. Most research has been done in the area of management by exception (MBE) reporting ([JPW81]). Such reporting is supposed to cause corrective responses to significant deviations from the norm, while eliminating the time consuming review of insignificant variations. If the analyst makes his selection criteria too tight, important deviations might be missed, so called "error of omission", whereas if he does not discriminate enough this may lead to an "alarm neurosis" creating too much detail.

Judd et al. [JPW81] report that most MBE systems use one of two criteria: percentage difference and absolute difference. Using percentage variance, the analyst selects all items that exceed a percentage threshold, e.g., a list of all customer accounts that have exceeded their credit limit by more than 10 %. Using absolute variance, the analyst selects all items that exceed a quantity threshold, e.g. a list of all customers whose accounts are 60 days past due.

In a study on budget exception reporting among 116 practicing managers Judd et al. [JPW81] found that the measure $p \times a$ (p = percentage variance, a = absolute variance)

slot name	slot entry
variable	VARIABLE NAME
normtype	NORM TYPE
function	FUNCTION DESCRIPTION
lowbound	LOW THRESHOLD
upbound	HIGH THRESHOLD

Table 3.2: norm unit

could best discriminate between significant and insignificant deviations. This combined measure reflects the phenomenon that as the base figures went up, subjects were more sensitive to smaller percentages; as the base figures went down, a large percentage was required for the deviation to be selected as significant.

From the preceding discussion it is clear that selection of the proper reference object, and selection of the variable(s) for which a comparison between a company and the reference object should be made, is fairly situation dependent. Therefore we have chosen to make the representation of the norm model as general as possible, and to allow the model builder to specify the parameters of this model. In table 3.2 the general structure of a norm unit is given. For an example of a norm unit, the reader is referred to table 4.12. The "variable" slot specifies for which variable a norm is being given. This will typically be a key performance indicator such as profit margin or return on assets. The "normtype" slot states which reference object serves as object of comparison with respect to that variable. Typical instantiations for this slot are "industry average", "previous year value", and "budget", indicating that the norm for the actual value of a variable is its industry average, previous year value, and budget value respectively. The "function" slot specifies which function (f(actual value, norm value)) is used as a measure of the difference between actual value and norm value, e.g. percentage difference. Finally, the "lowbound" and "upbound" slots specify the lower function value and upper function value, below and above which the difference between norm and actual value is considered to be significant.

3.3 The business model

The business model M represents relevant financial and operating variables and the relations among them. We distinguish between a) quantitative and b) qualitative relations among model variables:

a) $Y = f(X_1, \ldots, X_n)$.

b) $Y \leftarrow \{(X_1, sign_1), \ldots, (X_m, sign_m)\}$, where $sign_i \in \{\text{pos,neg}\}$.

For an example of a business model, the reader is referred to the IFC model on page 68. Quantitative model relations are not required to be additive. As we have already mentioned in section 1.4.4, non-additive equations are quite common in financial models. However, we do impose some restrictions on the data set that is used as input for the explanation process. These restrictions will be discussed in section 4.2.4. Quantitative relations are for example used to represent definitions, such as $profit = revenues - costs$, or consolidation equations. Quantitative relations can also be used to represent causal relations, determined by econometric analysis. In most cases, however, knowledge of causal relations will be available only in qualitative form. The second type of relation expresses knowledge where the exact quantitative relation among variables is not known. For example, "an increase in competition will lower the selling price" is represented as follows:

selling price $\leftarrow$ {(competition, neg)}.

The relevance of qualitative knowledge for economic analysis has been confirmed by a number of studies ([Bou78, Hux88, Ber92]). Such qualitative relations are bipolar, i.e. an increase in competition leads *ceteris paribus* to a price decrease, and a decrease in competition leads *ceteris paribus* to an increase in price. It is important to note here that the business model has a "mixed ontology". Quantitative functional relations refer to a state of the system at a particular point in time, whereas qualitative relations refer to *differences* between states.

The form of model relations is such that exactly one variable appears on the left hand side of a relation. This variable should be either:

- defined in terms of the variables appearing on the right hand side,

- a consolidation of variables appearing on the right hand side, or

- causally influenced by variables appearing on the right hand side.

Furthermore, no variable appears on the left-hand side of more than one relation. Economists often specify their models in such a way that the above requirements are met. Boutillier ([Bou84]) calls this the *reading* of the model. Rewriting an equation to get a different left hand side variable may not change the solution to the equation, but it does change its reading. The reading of the model is important for explanation generation and diagnostic reasoning; therefore relations should be represented according to their proper reading.

Time is represented by finite fixed-length time periods. This is customary for business models since company related data are usually collected for fixed time periods, e.g. weeks, months, and years. It is allowed that variables on the right-hand side of a relation refer to previous time periods. This might reflect a situation where there is a time lag between cause and effect. However, for ease of exposition we will assume that all variables refer to the same time period, since this makes no fundamental difference for the results of this thesis.

With the business model M we associate a directed graph $E(M) = (V, E)$, called the *explanatory* graph of M, as follows. The vertex set V contains as elements all variables appearing in the model. The edge set E contains a directed edge from vertex x_i to x_j iff:

1. $X_j = f(\ldots, X_i, \ldots) \in M$, or

2. $X_j \leftarrow \{\ldots, (X_i, sign_i), \ldots\} \in M$.

For an example of an explanatory graph, see figure 4.1. A restriction we impose on the model M is that its explanatory graph $E(M)$ should not contain any cycles, since this would make explanations circular. This restriction excludes models that contain simultaneous equations. Simultaneous equations usually arise because the model has a time structure that is too coarse-grained to represent causalities adequately. Events that in reality take place sequentially, are assumed to occur simultaneously at one moment in time. Theoretically this problem can be removed by taking a time structure that is less coarse-grained, in order to identify the time lag between cause and effect. In practice it may not be possible to obtain the detailed data required to model the time lag between cause and effect.

Nodes in the explanatory graph, with zero indegree, represent variables that can not be explained in M.

In the next chapter we shall provide an operational definition of the concept of explanation, using both the quantitative and qualitative relations of the business model. This will lead to the development of novel reasoning methods for the diagnosis of business performance.

Chapter 4

A formal framework for diagnosis and explanation

4.1 Introduction

In the previous chapter we presented the knowledge structures that are essential to model based diagnosis of business performance. We distinguished between a norm model, which specifies "normal" performance for particular key-variables, and a business model, which describes qualitative and quantitative relations among financial and operational variables.

In this chapter we develop new methods of model based diagnosis and explanation in the business domain. We build upon the approaches to diagnosis and explanation discussed in chapter 1, making improvements and extensions where appropriate.

In the conclusions of chapter 1 we argued that although a situation of incomplete information is almost a defining characteristic of diagnosis in the medical and technical domains, this is certainly not the case in the business domain. On the other hand, situations of incomplete information do occur in the latter domain, and our theory should be able to cope with them. Therefore we develop diagnostic methods for the following situations:

1. There is *complete* information about the actual values and norm values of variables in the business model.

2. There is *incomplete* information about the actual values and norm values of variables in the business model.

In *both* cases, explanations may be supported by qualitative relations. Qualitative relations represent *imprecise* rather than *incomplete* information about the strength of causal relations. We use the term "incomplete information", to refer to incomplete information concerning the actual and norm values of model variables.

In the situation where all actual and norm values are known, the problem of diagnosis and explanation reduces to *selecting* relevant influences from the available information. This problem has been addressed by Kosy and Wise ([Kos89, KW84]) and Courtney et al.([CPAM87]), whose work we discussed in sections 1.4.1 and 1.4.2 respectively. In section 4.2 we present a new method for diagnosis and explanation with complete information. The method of Kosy and Wise for generating explanations will prove to be an acceptable starting point, but it has some serious shortcomings. We shall make substantial changes in order to remove these shortcomings, and obtain the properties of aleatory explanations which we discussed in section 2.3.

If not all actual and norm values are known, i.e., in the case of incomplete information, the problem of explanation and diagnosis is one of finding consistent hypotheses. This problem has been addressed, be it in a rather informal manner, by Bouwman ([Bou78, Bou83]), whose work we discussed in section 1.4.3. In section 4.3 we present our approach to diagnosis and explanation with incomplete information, inspired by Bouwman's work. However, our approach will be given a sound formal basis by using concepts from the theory of qualitative reasoning.

4.2 Diagnosis and explanation with complete information

In sections 1.4.1 and 1.4.2 we discussed two approaches to diagnosis and explanation with complete information in the business domain. In section 1.4.4 we concluded that the method of Kosy and Wise has some advantages compared with the structural modeling approach of Courtney et al.:

1. The ability to generate explanations based on non-additive functions.

2. The automatic selection of significant influences.

We consider both properties to be desirable for diagnostic systems in the business domain. The ability to generate explanations based on non-additive functions is crucial because many financial relations are expressed by non-additive functions, for example, financial ratios and other definitions that relate financial and operational variables. The automatic selection of *significant* influences is vital in order to avoid "information overload", which would be created by including all influences, whatever their size.

On the other hand, we argue that the realisation of these properties in Kosy and Wise's system, has a number of shortcomings. These shortcomings will be analysed in section 4.2.1. Based on this analysis, we describe a new method of explanation in section 4.2.2. One extension of this new method is its ability to generate explanations based on qualitative relations in the business model.

Using this new method of explanation, we define the concept of diagnosis and illustrate both explanation and diagnosis with examples from the interfirm comparison model described in section 4.2.3. Finally, in section 4.2.4 we define the conjunctiveness constraint. If the actual and norm values involved in an explanation satisfy this constraint, our method of explanation yields valid results.

4.2.1 Analysis of Kosy and Wise's approach

In this section we analyse the method of explanation presented by Kosy and Wise in [KW84] and Kosy in [Kos89]. Since we have already presented this method in section 1.4.1, it suffices to point out its shortcomings in this section.

Firstly, we return to the distinction, made in section 2.3, between the explanation of a *quantitative* difference and the explanation of a *qualitative* difference. We stated there that in case the objective is to explain a *quantitative* difference, the set of counteracting causes is always empty. More specifically:

1. If the explanation of a *quantitative* difference, denoted by ΔY, is required, then all right-hand side variables in the equation for Y that have different values in the

XYZ	1990	1991	ε
product X	200	150	-50
product Y	200	245	$+45$
product Z	200	155	-45
total sales	600	550	

Table 4.1: Sales per product of XYZ-company

situations that are compared, are considered to be *contributing* causes.

2. If the explanation of a *qualitative* difference, denoted by ∂Y, is required, then all right-hand side variables in the equation for Y that have different values in the situations compared, are either contributing or counteracting causes, depending on the directions of their influences.

The direction of influence of a single variable is indicated in Kosy and Wise's system by its ε value. For any method of explanation that has been designed to explain differences, it should be clear whether it is intended to explain quantitative or qualitative differences.

The procedure as discussed in [Kos89] is intended to explain the *direction of difference*, or in our terminology the qualitative difference. Therefore it is surprising that the adequacy of an explanation is determined by looking at the fraction of the quantitative difference ΔY that it explains. The following example illustrates how this can yield counterintuitive explanations. Consider the data provided in table 4.1. This table contains the sales results of the XYZ-company for two consecutive years, for three different products sold by the company. Hence the corresponding equation in the business model would be:

$$\text{total sales} = \text{product X} + \text{product Y} + \text{product Z}$$

The objective is to explain the decline in total sales between 1990 and 1991. Application of the definition given in [Kos89], described in section 1.4.1, leads to the following result. The variables in $S^* = \{\text{product X, product Y, product Z}\}$ are added to the set of explanatory variables A in order of largest absolute effect (i.e. $|\varepsilon(\{X_i\}, Y)|$), until

$$1/\theta > \frac{\varepsilon(A, Y)}{\Delta Y} > \theta,$$

	t	$t+1$	ε
X	160	100	-60
V	10	-8	-180
W	-12	10	-176
Y	40	20	

Table 4.2: Data for non-conjunctive explanation

where θ is a number smaller than and close to 1. In the "sales" example the variable product X is added to A first, and the procedure will stop at that point, assuming $0 < \theta < 1$, with $A = \{\text{product X}\}$, since:

$$\frac{\varepsilon(\{\text{product X}\}, \text{total sales})}{\Delta\text{total sales}} = 1.$$

Apparently, another large negative influence on total sales, product Z, is not mentioned in the explanation. This is due to the effect of the increase in product Y, which compensated the decrease in product Z. This counterintuitive result is caused by two shortcomings in the concept of explanation used by Kosy and Wise:

- There is no a priori distinction between contributing and counteracting causes.
- The adequacy of an explanation is measured by the fraction of the quantitative difference ΔY that it explains, whereas the objective is to explain the qualitative difference ∂Y.

Secondly, it should be noted that the explanation procedure yields valid results only under specific restrictions. It is implicitely assumed that explanations are *conjunctive*. We illustrate this concept by an example; a formal definition is given in section 4.2.4. Consider the data provided in table 4.2, based on the equation

$$Y = X + (V \times W).$$

The objective is to explain the decrease in Y from time period t to $t+1$. The last column of table 4.2 shows the individual influences (ε) of the variables appearing in the equation for Y. Note that the influences of V and W taken separately are both negative. The influence of V and W *together* however is positive, i.e. $\varepsilon(\{V,W\},Y) = +40$. Hence, two contributing causes to the decrease in Y, turn into counteracting causes when considered simultaneously. Such non-conjunctive causes may not occur very often in practical diagnostic situations, but it should be stated explicitly that the method may not produce valid explanations for such situations.

4.2.2 A new model of explanation and diagnosis

In this section we present a new model of explanation and diagnosis in the business domain. This model has evolved considerably over a period of time; earlier versions can be found in ([DF92, Fee92, DF91]).

Recall from section 2.3 the canonical format for aleatory explanations:

$< a, F, r >$ because Φ, despite Ψ,

where $< a, F, r >$ specifies the "event" that object a has property F, whereas reference object r does not.

In this section we develop a method of explanation for the diagnosis of business performance that is based on this canonical format. We are interested in explaining the difference between the actual and norm behaviour of a particular business company. Consequently we have to explain the following type of events:

- $a =$ the actual behaviour of a company;
- $F =$ a particular variable deviates from its norm value;
- $r =$ the norm behaviour for the company involved.

Since the object a and reference object r will always be clear from the context, we simplify the *canonical explanation* format to:

$\partial Y = q$ occurred because C_b, despite C_a.

In this expression, $\partial Y = q$ plays the role of event, and states the qualitative difference between the *actual* and *norm* values of Y, denoted by Y^{act} and Y^{norm} respectively. This qualitative difference can take on one of the values {low, normal, high}, and is

	∂Y
$Y^{act} > Y^{norm}$	high
$Y^{act} = Y^{norm}$	normal
$Y^{act} < Y^{norm}$	low

Table 4.3: Mapping to qualitative value

determined according to the rules of table 4.3. For the purpose of diagnosis we are not interested in explaining $\partial Y =$ normal, since it is only required to explain why a variable deviates from its norm value. Note furthermore that an explanation for $\partial Y =$ normal would actually be an explanation for a quantitative difference, i.e. it would explain why $\Delta Y = Y^{act} - Y^{norm} = 0$.

The set of *contributing* causes is denoted by C_b and the set of *counteracting* causes by C_a. The way in which contributing and counteracting causes are determined depends on the type of relation (quantitative or qualitative) in the business model that sustains the explanation.

First, we discuss the situation where the explanation is sustained by a quantitative equation from the business model:

$$Y = f(X_1, \ldots, X_n).$$

To determine the contributing causes (C_b) and counteracting causes (C_a) that explain the qualitative difference between the actual and norm value of Y, i.e., the event $\partial Y = q$, we define a measure of influence as follows:

$$\inf(X_i, Y) = f(X_1^{norm}, \ldots, X_{i-1}^{norm}, X_i^{act}, X_{i+1}^{norm}, \ldots, X_n^{norm}) - Y^{norm}.$$

In words, $\inf(X_i, Y)$ gives the difference between the actual and norm values of Y if *only* X_i would have deviated from its norm value. Determination of the influence of X_i on Y involves the evaluation of a hypothetical situation represented by the expression $f(X_1^{norm}, \ldots, X_{i-1}^{norm}, X_i^{act}, X_{i+1}^{norm}, \ldots, X_n^{norm})$. Our "inf-measure" is actually the "mirror image" of Kosy and Wise's ε measure (see section 1.4.1). The inf-measure as such does not have any substantial advantages; we defined it in this way to emphasize the fact that our concept of explanation is substantially different, and to maintain a clear analogy with explanations based on qualitative relations that will be discussed later in this section.

The correct interpretation of the inf-measure depends on the form of the function f. For the moment we assume that previous period values serve as norm values, so we are effectively explaining a change of Y between periods $t-1$ and t. If f is additive, then $\inf(X_i, Y)$ is correctly interpreted as a quantitative specification of the change in Y that is explained by the change in X_i from $X_{i,t-1}$ to $X_{i,t}$. If, however, f is non-additive, then the interpretation of $\inf(X_i, Y)$ is considerably more difficult. As we stated in section 2.3, it can certainly *not* be interpreted as a quantitative specification of the change in Y that is explained by the change in X_i alone, since the value of $\inf(X_i, Y)$ also depends on the $t-1$-level at which other variables (besides X_i) are held constant. Hence, such a quantitative claim only holds true within that particular context.

We define contributing and counteracting causes as follows:

Definition 4.1 (Contributing Causes) *The set C_b of causes contributing to $\partial Y = q$ consists of the variables $X_i \in \{X_1, \ldots, X_n\}$ such that inf(X_i, Y) $\times \Delta Y > 0$.*

Definition 4.2 (Counteracting Causes) *The set C_a of causes counteracting $\partial Y = q$ consists of the variables $X_i \in \{X_1, \ldots, X_n\}$ such that inf(X_i, Y) $\times \Delta Y < 0$.*

In words, the contributing causes are those variables whose influence values have the same sign as ΔY, and the counteracting causes are those variables whose influence values have the opposite sign.

We apply these definitions to the "sales" example that was presented in table 4.1. A full specification of the event to be explained is:

$<$ XYZ (1991), ∂total sales = low, XYZ (1990) $>$,

since the objective is to explain a decrease in total sales from 1990 to 1991 for the XYZ-company. Computation of the influences of the individual variables in the equation for total sales yields the following results:

ABC-company	1990	1991	inf
sales volume (sv)	100000	150000	+50000
variable sales expenses (vse)	1	1.05	+5000
fixed sales expenses (fse)	150000	180000	30000
total sales expenses (tse)	250000	337500	

Table 4.4: Data for ABC-example

inf(product X, total sales) $= -50$,

inf(product Y, total sales) $= 45$,

inf(product Z, total sales) $= -45$.

Hence the following explanation is obtained:

∂total sales = low, because C_b = {product X, product Z},
despite C_a = {product Y}.

Note that one of the problems that we signalled in section 4.2.1, namely the omission of significant influences from an explanation, has been removed simply by conforming to the format of aleatory explanations, i.e., by making a strict a *priori* separation between contributing and counteracting causes. The explanation now mentions the significant contributing cause "product Z" that was omitted before.

As currently defined, our explanation method does not possess the desired property of leaving insignificant influences out of the explanation. This could lead to "information overload". It is therefore not surprising that observation of human financial analysts shows that they tend to leave out insignificant influences from the explanation.

In order to illustrate this point, we consider the example in table 4.4. This example involves the increase of total sales expenses (tse) between 1990 and 1991 for the ABC-company. An exact specification of the event to be explained in the format $< a, F, r >$ yields:

<ABC(1991), ∂tse = high, ABC(1990)>

Thus the behaviour of the ABC-company in 1991 specifies the actual behaviour, and the behaviour of the ABC-company in 1990 specifies norm behaviour. The equation in the business model that corresponds to this example is:

tse = vse × sv + fse

Application of the definitions for contributing and counteracting causes yields the following explanation of this event:

∂tse = high because C_b = {sv, vse, fse}, despite C_a = {}.

Since all variables contributed to the increase in "total sales expenses", the set of counteracting causes is empty. On the other hand, it seems rather superfluous to include "variable sales expenses" as a contributing cause, since it only explains a relatively small fraction of the increase. In order to leave the insignificant influence of "variable sales expenses" out of the explanation, we first define the influence of a set of variables, which is a straightforward extension of the influence of a single variable.

Definition 4.3 (Influence of set of variables) *The influence of a set of variables $A \subseteq \{X_1, \ldots, X_n\}$ on Y is defined as $inf\,(A, Y) = f(A^{act}, \bar{A}^{norm}) - Y^{norm}$.*

Here $\bar{A}$ denotes the complement of A, $f(A^{act}, \bar{A}^{norm})$ denotes the value that results from replacing all variables in A by their actual value, and all variables in $\{X_1, \ldots, X_n\} - A$ by their respective norm values.

Definition 4.4 (Parsimonious set of contributing causes) *Given T_b, the parsimonious set of contributing causes C_b^p is the smallest subset of C_b such that $\frac{inf(C_b^p, Y)}{inf(C_b, Y)} \geq T_b$.*

Definition 4.5 (Parsimonious set of counteracting causes) *Given T_a, the parsimonious set of counteracting causes C_a^p is the smallest subset of C_a such that $\frac{inf(C_a^p, Y)}{inf(C_a, Y)} \geq T_a$.*

The parsimonious set of contributing causes is the smallest subset of the set of contributing causes such that its influence on Y exceeds a particular fraction (T_b) of the influence of the complete set. In case there are several sets of equal cardinality that explain a

∂Y	inv(∂Y)
high	low
normal	normal
low	high

Table 4.5: definition of inv-operator

$\otimes$	low	normal	high
pos	low	normal	high
neg	high	normal	low

Table 4.6: Definition of $\otimes$

fraction larger than T_b, the one with the highest inf-value is called the parsimonious set. The definition with respect to counteracting causes is clearly analogous.

The fractions T_b and T_a are numbers between 0 and 1, and will usually be close to 1. We speak of a parsimonious one-level explanation:

$\partial Y = q$ because C_b^p, despite C_a^p.

Let us apply this new definition, with $T_b = 0.8$, to the example of table 4.4. We now obtain the following sets of parsimonious contributing and counteracting causes:

$$C_b^p = \{\text{sv, fse}\}$$

$$C_a^p = \{\}$$

Note that taking the parsimonious set of contributing causes rather than the complete set, removes the relatively insignificant influence of "variable sales expenses" from the explanation.

Thus far we have been concerned with explanations based on quantitative equations in the business model. Now we direct our attention to explanations based on qualitative relations. Recall from chapter 3 that a qualitative relation only states whether a variable has a positive (pos) or negative (neg) influence on another variable. Since the strength of such an influence is not quantified, we only look at the *qualitative* difference between actual and norm value when determining contributing and counteracting causes.

Definition 4.6 (Contributing causes, qualitative relations) *The set of contributing causes for $\partial Y = q$ in case $Y \leftarrow \{(X_1, s_1), \ldots, (X_m, s_m)\} \in M$, is the set $\{X_i | \partial X_i \otimes s_i = \partial Y, i \in \{1, 2, \ldots, m\}\}$.*

Definition 4.7 (Counteracting causes, qualitative relations) *The set of counteracting causes for $\partial Y = q$ in case $Y \leftarrow \{(X_1, s_1), \ldots, (X_m, s_m)\} \in M$, is the set $\{X_i | \partial X_i \otimes s_i = inv(\partial Y), i \in \{1, 2, \ldots, m\}\}$.*

Where $s_i \in$ {pos, neg}; the "inv" operator is defined in table 4.5, and the $\otimes$ operator is defined in table 4.6. For example, if X has a positive influence on Y, and $\partial X =$ low, and $\partial Y =$ low, then X is a contributing cause of $\partial Y =$ low. On the other hand, if X would have a negative influence, then X would be a counteracting cause.

In case of explanations sustained by qualitative relations the distinction between complete and parsimonious explanations is no longer possible, since there is no possibility to weigh the different contributing and counteracting causes. Consequently, $C_b^p = C_b$ and $C_a^p = C_a$.

Thus far, we have discussed "one-level" explanations. These explanations are one level deep, in the sense that they are based on a single relation from the business model. For diagnostic purposes, however, it is useful to continue an explanation of $\partial Y =$ q, by explaining the qualitative differences between the actual and norm values of its contributing causes. This process can be continued until a contributing cause is encountered that cannot be explained within the business model, because the business model does not contain a relation in which this contributing cause appears on the left-hand side. This idea is clearly similar to the "causation tree" that we encountered in Courtney et al.'s work (see section 1.4.2), and the "tree of causes" encountered in Bouwman's work (see section 1.4.3).

More generally, we define a maximal explanation for $\partial Y = q$ as follows.

Definition 4.8 (Maximal explanation) *A maximal explanation for $\partial Y = q$ is a tree with the following properties:*

1. *Y is the root node of the tree.*

2. *the root node has two types of children, corresponding to its parsimonious contributing and counteracting causes respectively.*

3. *a node that corresponds to a contributing cause has two types of children, corresponding to its parsimonious contributing and counteracting causes respectively.*

4. *a node that corresponds to a counteracting cause has no children.*

5. *a node that corresponds to a variable that cannot be explained in the business model has no children.*

For an example of a maximal explanation, the reader is referred to figure 4.2. This definition clearly reflects the different roles of contributing and counteracting causes. Without contributing causes there would be no explanation at all, whereas the absence of counteracting causes merely indicates that there was no opposition against $\partial Y = q$. Therefore, a maximal explanation for $\partial Y = q$ continues with its contributing causes, whereas the counteracting causes are not explained any further.

A diagnosis is an explanation for observed abnormal behaviour of a company. Before we proceed to the final definition of this section, namely the definition of a diagnosis, we return to the detection of abnormal behaviour, also called *problem identification*. We have already touched upon this subject in section 3.2, when the norm model was discussed. Suppose the norm model contains a norm unit for variable Y. If the function-slot of the norm unit for Y contains the function g, then we define:

$$\text{Deviation} = g(Y^{act}, Y^{norm})$$

Given lowbound(Y) and upbound(Y), there are three possibilities with respect to the value of Deviation:

1. Deviation $<$ lowbound(Y), in this case the symptom $\partial Y = \text{low}$ has been discovered.

2. Deviation $>$ upbound(Y), in this case the symptom $\partial Y = \text{high}$ has been discovered.

3. lowbound(Y) $\leq$ Deviation $\leq$ upbound(Y), in this case no symptom has been discovered,

where lowbound(Y) and upbound(Y) denote respectively the value in the lowbound and upbound-slot of the norm-unit for Y. The result of problem identification is a set of symptoms $S = \{\partial Y_1 = q_1, \ldots, \partial Y_n = q_n\}$[1], where $q_i \in$ {low,high}. We define the data set I as the set that contains all actual values and norm values for the variables in M.

Definition 4.9 (Diagnosis for $< M, I, S >$) *A diagnosis for $< M, I, S >$ is the set of maximal explanations in business model M for all symptoms in S, using data set I.*

Hence, a diagnosis provides maximal explanations for all observed symptoms. Since the norm-type applied for the identification of a symptom determines the norm-type that is applied to its causal influences, the data set I may contain several norm behaviours, besides actual behaviour,.

In the next section we illustrate our theory of explanation through a textbook example that is based on a case study.

4.2.3 Illustration: an interfirm comparison model

The model that we present in this section has been taken from the book "Interfirm comparison" by Herbert Ingham and L. Taylor Harrington ([ITH80]). The model was originally developed for a case study, in order to demonstrate how management can use interfirm comparison (IFC) to diagnose the strengths and weaknesses of its business, and take remedial action. This case study involved the comparison of nine firms that operate in a section of the mechanical engineering industry. The firms in this particular section manufacture a class of industrial products needed by their customers as components of their own products.

Translation of the interfirm comparison model into the formalism for business models that we presented in section 3.3, yields the following equations and qualitative relations:

1. ROA = PM $\times$ TA

[1]Strictly speaking, a symptom is a triple $< a, \partial Y = q, r >$, but we save notation by assuming that a and r are clear from the context

2. $\text{PM} = 1 - (\text{AC} + \text{PC} + \text{RDC} + \text{DMC})$

3. $\text{TA} = 1 \div \text{OA}$

4. $\text{OA} = \text{CA} + \text{FA}$

5. $\text{PC} = \text{MC} + \text{WLC} + \text{WS} + \text{OPC}$

6. $\text{CA} = \text{FGS} + \text{DEB} + \text{MS} + \text{WIP}$

7. $\text{FA} = \text{LB} + \text{PAM} + \text{OFA}$

8. $\text{WIP} \leftarrow \{(\text{SPP}, \text{neg})\}$

9. $\text{MC} \leftarrow \{(\text{MU}, \text{pos}), (\text{BC}, \text{pos})\}$

10. $\text{DMC} \leftarrow \{(\text{SPP}, \text{pos}), (\text{EX}, \text{pos})\}$

11. $\text{FGS} \leftarrow \{(\text{SPP}, \text{pos})\}$

12. $\text{DEB} \leftarrow \{(\text{EX}, \text{pos})\}$

In the remainder of this section we will explain the variables and relations that appear in the model. Return on operating assets (ROA) gives an overall indication of how profitably a company's management is using the resources available to it. This variable is split up into profit margin (PM) and turnover of assets (TA) (equation 1). Profit margin shows what profit has been earned on sales and turnover of assets shows how intensively the firm uses the available assets. The firm's profit margin on sales is determined by its departmental cost ratios (administrative costs (AC), production costs (PC), research & development costs (RDC) and distribution and marketing costs (DMC)), which express the cost falling under four headings as a percentage of sales (equation 2). The higher the cost ratios, the lower the profit margin on sales. The variables work sub-contracted (WS), materials costs (MC), works labour costs (WLC) and other production costs (OPC) provide a breakdown of production costs (PC) (equation 5). Interfirm comparison of these ratios, which express the production cost incurred under four headings as fractions of

Type	sales/production policy
A	predominantly quantity production of standard products made for stock in anticipation of order
B	predominantly quantity production to customers' requirements after receipt of orders
C	predominantly small quantity production to customers' orders, with occasional small batch production for stock

Table 4.7: Sales and production policy

the sales value of goods produced, will show how far each of the main areas of production cost accounts for differences between a firm's production cost to sales and those of others.

A firm whose cost of bought-out components is high in relation to its total material cost, will tend to have a relatively high materials cost ratio (MC), because the cost of these parts will include the production and other costs incurred and the profit margin taken by the suppliers (relation 9). The materials-cost ratio is also affected by the kind of metal, namely ferrous or non-ferrous, that the company uses in its production process (MU). Non-ferrous metals are generally more expensive. The firm's distribution and marketing cost ratio (DMC) is mainly influenced by its fraction of export sales (EX) and its sales/production policy (SPP) (relation 10). A high fraction of export sales tends to lead to higher distribution and marketing costs. Also a firm that uses sales/production policy type A, i.e. predominantly quantity production of standard products made for stock in anticipation of orders, will generally have higher distribution and marketing costs than firms that use policies B or C (see table 4.7).

Turnover of assets (TA) is a simple transformation of return on assets (ROA) and indicates the asset utilization rate of a firm (equation 3). The lower this figure, the higher the rate of asset utilization. This asset utilization rate is sub-divided into the utilization rate of current assets (CA) and fixed assets (FA) (equation 4). The variables work in progress (WIP), finished goods stock (FGS), materials stock (MS), and debtors (DEB) provide a breakdown of the current assets ratio (equation 6). Interfirm comparison of WIP, FGS, MS, and DEB shows to what extent each of the main components of current assets accounts for interfirm differences in CA. The variables land and buildings (LB), plant and machinery (PAM), and other fixed assets (OFA) give a breakdown of the fixed-assets ratio (equation 7). The ratio of debt to sales (DEB) is influenced by the fraction of export sales, since longer credit terms may have to be given to foreign customers (relation 12).

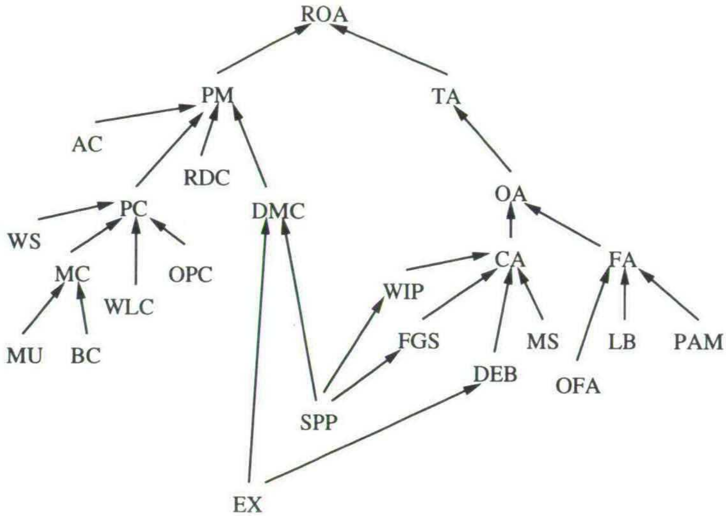

Figure 4.1: Explanatory graph E(IFC) for IFC model

The graph in figure 4.1 shows the explanatory graph $E(IFC)$, as defined in section 3.3, for the interfirm comparison model.

Most of the variables take their value from the real numbers, except for SPP (sales/production policy) and MU (metal used). Note that the latter variables appear only in qualitative relations. Sales/production policy can have the values A, B, and C (see table 4.7), which means that it is measured on a nominal scale[2]. With respect to the relations between sales/production policy and the other variables, this scale is converted to an ordinal scale, i.e., the basic operation on its elements is determination of greater or less. In this model, the values A, B, and C are ordered: $A > B = C$. The variable MU stands for the metal that is predominantly used in the production process. This variable can take on two values: predominantly ferrous and predominantly non-ferrous. Regarding the relation between MU and materials cost/sales value of production (MC),

[2]The only basic operation allowed on a nominal variable is determination of equality

the values are ordered non-ferrous > ferrous, reflecting the fact that non-ferrous metals are more expensive.

Table 4.8 gives the data of nine different firms operating in the mechanical engineering industry, for all variables that appear in the IFC-model. The data of these firms have been taken from [ITH80], with the additional note however that the values of the variables SPP (A, B and C) and MU (ferrous and non-ferrous), have been replaced by "order preserving" integer values, in order to enable the computation of an industry average. The column "average" has been added to the original data to represent norm behaviour in the examples that follow. The average values have been determined as follows:

- For all variables that do not appear on the left hand side of an equation, the average has been computed directly by taking the mean value of all nine firms.
- For each equation $Y = f(X_1, \ldots, X_n)$, Y^{avg} has been computed by taking $Y^{avg} = f(X_1^{avg}, \ldots, X_n^{avg})$.

This procedure has been followed in order to guarantee that the "average state" is internally consistent, which would not be the case if all averages would have been computed directly by taking the mean value of the nine firms.

Apart from being an illustration of our theory, the example is at the same time a *test* of the theory, since we will compare the explanations provided in the book from which this case was taken ([ITH80]) with the results obtained by our theory.

In ([ITH80] p.35) the following explanation is given for firm1's relatively, i.e. compared to other firms, high return on operating assets (ROA):

> *Why is firm1's return on operating assets relatively high ? Comparisons of ratios 2 (PM) and 3 (TA) with those of the other firms show that the firm's high ratio 1 (ROA) is due to a combination of a comparatively high profit on sales and a comparatively fast turnover of assets.*

We specify the event to be explained as $<$ firm 1, ∂ROA = high, industry average $>$. In the quoted text, industry average is not mentioned explicitly as a reference object. It is stated, however, that the values of PM and TA for firm 1, are compared with those of the other firms. Therefore it is justified to take the industry average as the reference

variables	FIRMS									
	1	2	3	4	5	6	7	8	9	avg
ROA	0.251	0.239	0.189	0.161	0.133	0.132	0.088	0.079	0.041	0.137
PM	0.19	0.199	0.151	0.099	0.103	0.115	0.087	0.089	0.047	0.12
TA	1.32	1.2	1.25	1.63	1.29	1.15	1.01	0.9	0.87	1.14
OA	0.758	0.833	0.8	0.613	0.775	0.869	0.99	1.111	1.149	0.877
PC	0.628	0.635	0.711	0.747	0.725	0.719	0.754	0.774	0.802	0.722
RDC	0.005	0.01	0.009	0.007	0.007	0.011	0.014	0.0	0.002	0.007
DMC	0.109	0.116	0.047	0.072	0.062	0.058	0.073	0.046	0.064	0.072
AC	0.068	0.04	0.082	0.075	0.103	0.097	0.072	0.091	0.085	0.079
MC	0.32	0.287	0.339	0.301	0.397	0.316	0.336	0.347	0.358	0.333
WLC	0.165	0.221	0.232	0.283	0.157	0.241	0.248	0.274	0.289	0.234
OPC	0.143	0.127	0.14	0.092	0.104	0.113	0.114	0.153	0.155	0.127
WS	0.0	0.0	0.0	0.071	0.067	0.049	0.056	0.0	0.0	0.027
CA	0.465	0.481	0.412	0.35	0.369	0.449	0.549	0.582	0.608	0.474
FA	0.293	0.352	0.388	0.263	0.406	0.42	0.441	0.529	0.541	0.404
MS	0.08	0.11	0.081	0.079	0.068	0.092	0.1	0.101	0.082	0.088
WIP	0.043	0.04	0.063	0.062	0.083	0.106	0.188	0.225	0.245	0.117
FGS	0.132	0.102	0.047	0.037	0.039	0.045	0.044	0.053	0.057	0.062
DEB	0.21	0.229	0.221	0.172	0.179	0.206	0.217	0.203	0.224	0.207
LB	0.13	0.158	0.194	0.169	0.208	0.214	0.311	0.277	0.268	0.214
PAM	0.16	0.189	0.19	0.089	0.194	0.203	0.123	0.246	0.264	0.184
OFA	0.003	0.005	0.004	0.005	0.004	0.003	0.007	0.006	0.009	0.005
EX	0.16	0.4	0.15	0.0	0.25	0.23	0.1	0.12	0.3	0.19
SPP	2	2	1	1	1	1	1	1	1	1.22
BC	0.32	0.39	0.29	0.33	0.56	0.35	0.4	0.37	0.43	0.382
MU	2	1	2	1	2	1	1	2	1	1.44

Table 4.8: Data for interfirm comparison

Firm 1	norm	actual	inf
ROA	0.137	0.251	
PM	0.12	0.19	0.0796
TA	1.14	1.32	0.0214

Table 4.9: Data for explanation of ∂ROA = high

object, in order to make a comparison between the textbook analysis and the results of our model of explanation. Taking $T_b = T_a = 0.85$, the model yields the following results. In table 4.9 comparison is made between ROA of firm 1 and industry average (norm). From the data in table 4.9 it follows that $C_b^p = \{\text{PM,TA}\}$, since both profit margin and turnover of assets contribute to the difference between norm value and actual value, and both are needed to explain the desired fraction of inf(C_b, ROA). Obviously, $C_a^p = \{\}$.

Comparison of human analysis and the result of our explanation method shows some noticable similarities. Firstly, both the human and the model explain the relatively high return on operating assets (ROA) in terms of its right-hand side variables, profit on sales (PM) and turnover of assets (TA). The model consistently uses the industry average as a comparison to explain the relatively high return on operating assets of firm 1. The second sentence of the textbook analysis gives no clue about how exactly these "comparisons of ratios 2 and 3 with those of the other firms" are made. Both the human analyst and the model state that both profit on sales and turnover of assets had significant influences.

The analysis continues as follows, [ITH80] p. 35:

> *Why is the firm's profit on sales ratio relatively high ? The most immediate answer is provided by its cost/sales ratios... Firm1's profit on sales is high mainly because its production cost ratio is the lowest of all... On the other hand, its distribution and marketing cost ratio is almost the highest of all.*

Analogous to the previous example, the event to be explained is specified as $<$ firm 1, ∂PM = high, industry average $>$. Table 4.10 summarizes the model results for the explanation of firm1's relatively high profit on sales. From the data in table 4.10 it follows that $C_b^p = \{\text{PC}\}$ and $C_a^p = \{\text{DMC}\}$. Notice that neither the human analyst nor the model mention that firm1's administrative cost/sales ratio is below average. The reason is that its contribution to the overall contributing influence (inf(C_b,PM)= 0.107)

Firm 1	norm	actual	inf
PM	0.12	0.19	
AC	0.079	0.068	0.011
PC	0.722	0.628	0.094
RDC	0.007	0.005	0.002
DMC	0.072	0.109	−0.037

Table 4.10: Data for explanation of ∂PM = high

Firm 1	norm	actual	∂	*sign*	$\partial \otimes$ *sign*
DMC	0.072	0.109	high		
SPP	1.22	2	high	pos	high
EX	0.19	0.16	low	pos	low

Table 4.11: Data for explanation of ∂DMC = high

on profit on sales is negligible. The same reasoning holds for research and development cost. This shows that both the human analyst and the model tend to leave insignificant influences out of the explanation. The human analyst also mentions a counteracting influence, by indicating that firm1's marketing cost ratio is almost the highest of all firms considered. This counteracting influence is also noticed by the explanatory model, since $C_a^p = \{\text{DMC}\}$.

The foregoing examples both involved quantitative equations. The next example shows how qualitative relations are used to generate explanations. Again we make a comparison between textbook analyses and model results. On page 36 of [ITH80] we find the following analysis:

> *Reverting to firm 1,...this firm applies sales/production policy A; it follows that its relatively high distribution and marketing cost ratio can probably be regarded as being part and parcel of its sales/production policy and therefore a condition of its success. The other factor ...- a high export percentage - would not have caused firm 1's distribution and marketing cost ratio to be high, since its percentage of export sales is comparatively small.*

Again, industry average seems to be the proper reference object, i.e., the event to be explained is < firm 1, ∂ DMC = high, industry average > From table 4.11 it can be concluded that $C_b^p = \{\text{SPP}\}$ and $C_a^p = \{\text{EX}\}$. The method of explanation generation for qualitative relations is clearly reflected in the above textbook fragment. The percentage

slot name	slot entry
variable	ROA
normtype	industry average
function	(actual value − norm value)/norm value
lowbound	−0.5
upbound	none

Table 4.12: Norm model for diagnostic example

of export sales of firm1 cannot explain its relatively high distribution and marketing costs because its export percentage is actually *below* average and there is a positive relation between export sales and distribution and marketing costs. This is reflected by the fact that the model of explanation lists percentage of export sales (EX) as a counteracting influence. An explanation can be found, however, by looking at firm1's sales/production policy. This is exactly the explanation that the model gives, since SPP is listed as the only contributing influence.

The previous examples of different one-level explanations show that the results of our method correspond to textbook analyses. We now provide an example of a complete diagnosis. Firstly the applicable norm model is specified in table 4.12. The diagnosis will be performed for firm 9, which is doing particularly bad compared to industry average. Problem identification yields the set of symptoms $S = \{\partial\text{ROA} = \text{low}\}$ since the relative difference between norm value and actual value, $(0.041 - 0.137)/0.137 = -0.7$, is below the specified lowbound of -0.5.

The diagnosis starts with a one level explanation of ∂ ROA = low. Application of definitions 4.1 and 4.2 yields: $C_b^p = \{\text{PM,TA}\}$ and $C_a^p(\text{ROA}) = \{\}$. Figure 4.2 summarizes the results of the complete diagnostic process, where dotted lines indicate counteracting causes. Since there is only one symptom to be explained, the diagnosis contains only one maximal explanation. Thus, figure 4.2 actually depicts the maximal explanation, as specified in definition 4.8, for ∂ROA = low.

4.2.4 The conjunctiveness constraint

In section 4.2.1 we pointed out the problem of non-conjunctive explanations for Kosy and Wise's method. In order to guarantee that the definitions of parsimonious contributing

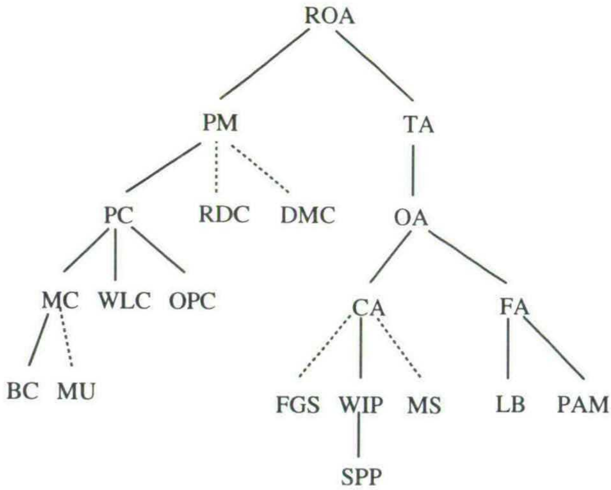

Figure 4.2: Diagnosis for $S = \{\partial\text{ROA} = \text{low at firm } 9\}$

and counteracting causes yield valid explanations, it is required that these causes are conjunctive. Therefore some constraints must be imposed on the situations for which an explanation can be given. The problem of non-conjunctive explanations could be solved by demanding that only additive functions are used to sustain explanations. This, however, would hinder the application to diagnosis and explanation in the financial domain, where non-additive functions are very common. Hence, rather than imposing restrictions on the functional form of the equations involved, restrictions are imposed on the data set, i.e. the actual values and norm values, that can be used in an explanation.

In order to define the conjunctivity property, the set $\{X_1, \ldots, X_n\}$ is partitioned into three sets:

Positive Influence Set (PI): the set of $X_i \in \{X_1, \ldots, X_n\}$ such that $\inf(X_i, Y) > 0$.

Negative Influence Set (NI): the set of $X_j \in \{X_1, \ldots, X_n\}$ such that $\inf(X_j, Y) < 0$.

No Influence Set (NOI): the set of $X_k \in \{X_1, \ldots, X_n\}$ such that $\inf(X_k, Y) = 0$.

Definition 4.10 (Conjunctiveness of data set) *A data set consisting of* $\{X_1^{act}, \ldots, X_n^{act}\}$ *and* $\{X_1^{norm}, \ldots, X_n^{norm}\}$ *is conjunctive with respect to the explanation of* ∂Y *iff for all* $i \in \{1, \ldots, n\}$*:*

	t	$t+1$	ε	inf
X	160	100	-60	-60
V	10	-8	-180	216
W	-12	10	-176	220
Y	40	20		

Table 4.13: Data for non-conjunctive explanation

1. $X_i \in PI$ *and* $A \subseteq \{X_1, \ldots, X_n\} - \{X_i\}$ *implies* $\mathit{inf}(A \cup \{X_i\}, Y) > \mathit{inf}(A, Y)$.

2. $X_i \in NI$ *and* $A \subseteq \{X_1, \ldots, X_n\} - \{X_i\}$ *implies* $\mathit{inf}(A \cup \{X_i\}, Y) < \mathit{inf}(A, Y)$.

3. $X_i \in NOI$ *and* $A \subseteq \{X_1, \ldots, X_n\} - \{X_i\}$ *implies* $\mathit{inf}(A \cup \{X_i\}, Y) = \mathit{inf}(A, Y)$.

This definition captures the intuitive notion that the influence of a variable does not "turn around" when it is considered in conjunction with a number of other variables. We will show that the example given in section 4.2.1 does not satisfy definition 4.10. For convenience we repeat table 4.2 in table 4.13, where an extra column has been added, containing the "inf" values of the individual variables.

In this example we have $PI = \{V, W\}$, $NI = \{X\}$ and $NOI = \{\}$, as can verified easily by looking at the fourth column of table 4.13. The conjunctiveness constraint is not satisfied since $V \in PI$ and $W \in \{V, W, X\} - \{V\}$ but $\text{inf}(\{V, W\}, Y) = 40$, which is significantly below the influence of W alone, i.e. $\text{inf}(W, Y) = 220$.

Since the data provided in table 4.13 do not satisfy the conjunctiveness constraints, our theory of explanation can not be applied in this situation.

4.3 Diagnosis and explanation with incomplete information

Diagnosis with incomplete information necessarily involves the formation of hypotheses. We already encountered a situation of incomplete information diagnosis in the business domain in the work of Bouwman, discussed in section 1.4.3. The solution proposed by Bouwman, based on extensive analysis of problem solving protocols, is to resort to qualitative characterisations of quantitative variables. In this way, each variable can take on only a fairly small number of alternative values, e.g., up, down, and stable; consequently the formation of hypothetical explanations in qualitative terms becomes tractable. In our approach to diagnosis with incomplete information, we follow the results of Bouwman's study, in the sense that reasoning is performed on qualitative characterisations of variables. We already used qualitative characterisations for diagnosis with complete information, namely in case explanations were supported by qualitative relations. In diagnosis with incomplete information, however, *all* reasoning is done qualitatively.

Unlike Bouwman, we are not interested in simulating human problem solving behaviour *in all respects*, i.e. including obvious shortcomings and limitations. Instead, we intend to model correct diagnostic reasoning with qualitative values. To this end, we make use of concepts from the area of Qualitative Reasoning, which provide a formal basis for qualitative diagnosis. These concepts are used to make a translation of the equations in the business model into qualitative relations. As a result, we obtain a transformed business model, denoted by M', that consists only of qualitative relations.

In the next section we describe the transformation of the business model using qualitative reasoning techniques. Thereafter, in section 4.3.2 we define the notions of explanation and diagnosis with incomplete information.

4.3.1 Qualitative Reasoning and model transformation

At the beginning of the 1980s, Qualitative Reasoning became a separate discipline within Artificial Intelligence, as witnessed by the appearance of a special issue in 1984 of the Artificial Intelligence Journal ([Be84]). The theory of Qualitative Reasoning has been proposed by several researchers ([DKB84, Kui86, For84]) as a way of incorporating deep knowledge into knowledge based (expert) systems. The field is also known as Qualitative Physics (QP), because the central domain of application is physics. Qualitative Reasoning techniques have also been employed in other domains such as medicine ([KK84]), electronic circuits ([Wil84]), and ecology ([Sim86]). Qualitative Physics is concerned with modeling the real world of time varying quantities in a symbolic, qualitative, and causal way. The main goals of Qualitative Physics are: 1) to be "simpler" than classical physics, yet retain all important distinctions in behaviour, 2) to produce causal accounts of physical mechanisms, and 3) to provide foundations for common-sense models for the next generation of expert systems ([DKB84]). More recently Qualitative Reasoning methods have been applied to the domain of economics, specifically macroeconomics ([Ber92, FL90]) and financial analysis ([Ham90, WC89, DF90, HBD86]). For an overview of different approaches to qualitative reasoning the reader is referred to [Iwa89].

A crucial issue in qualitative reasoning is the determination of the appropriate *quantity space* for variables, which is the set of qualitative values that a variable can have. In general, the appropriate quantity space is the simplest one that is able to represent all important distinctions in behaviour. Since we are concerned with explaining departures from "normality", the quantity space {low, normal, high} which we have already introduced in section 4.2, seems appropriate. Next we should define the semantics of relations among variables with respect to this quantity space. Table 4.14 defines the semantics of equations with respect to the quantity space {low, normal, high} and the operators $+, \times, -$, and $\div$. It is important to note that although equations represent laws that refer to the *values* of variables, they are used here to evaluate the qualitative *difference between values*. Therefore it is not surprising that the evaluation of equations with qualitative differences only yields valid outcomes under particular assumptions.

To give a simple example of the assumption underlying the definitions of the $\times$ and $\div$ operator: the expression *high* $\times$ *high* evaluates to *high* according to table 4.14. More specifically, when the norm is the value in the previous time period, it is inferred that the product of two increasing variables also increases. This inference, however, is not valid in general, but it is valid under the assumption that both the norm and the actual value of the multiplicands are positive. Recall that $\Delta Y = Y^{act} - Y^{norm}$. If $Y = A \times B \in M$, then we have

$$\Delta Y = A^{act} \times B^{act} - A^{norm} \times B^{norm},$$

and consequently $\Delta Y = B^{norm} \times \Delta A + A^{act} \times \Delta B$ or equivalently $\Delta Y = B^{act} \times \Delta A + A^{norm} \times \Delta B$. Thus, if either A^{act} and $B^{norm} > 0$ or A^{norm} and $B^{act} > 0$, then the qualitative differences ∂Y, ∂A, and ∂B satisfy

$$\partial Y \approx \partial A + \partial B,$$

$+, \times$	low	normal	high
low	low	low	?
normal	low	normal	high
high	?	high	high

$-, \div$	low	normal	high
low	?	low	low
normal	high	normal	low
high	high	high	?

Table 4.14: Definition of operators $+,\times,-,$ and $\div$

$\approx$	low	normal	high	?
low	true	false	false	true
normal	false	true	false	true
high	false	false	true	true
?	true	true	true	true

Table 4.15: Definition of weak equality

where weak equality ($\approx$) is defined in table 4.15. Therefore, qualitative addition and multiplication are equivalent if one assumes that only positive norm values and actual values occur.

Analogously, if $Y = A \div B \in M$, we get

$$\Delta Y = \frac{1}{B^{act}} \Delta A - \frac{A^{norm}}{B^{act} B^{norm}} \Delta B$$

Consequently, if B^{norm}, B^{act}, and $A^{norm} > 0$, then the qualitative differences ∂Y, ∂A, and ∂B satisfy

$$\partial Y \approx \partial A - \partial B,$$

which justifies the equivalence of qualitative division and subtraction.

These assumptions are not unduly restrictive in the financial domain, but earlier research ([Bou78, WC89]) failed to state them explicitly. For example, Bouwman's operator definitions for division and subtraction (see chapter 1, table 1.6), defined on the qualitative value set {down, stable, up}, are identical to our definitions in table 4.14.

For the purpose of qualitative explanation and diagnosis, the equations in business model M are translated into qualitative relations, resulting in a transformed model M'. However, in order to guarantee the validity of this translation process, certain restrictions must be imposed on the expression on the right-hand side of an equation. We call an expression that is allowed to occur on the right-hand side of an equation, a well formed expression, or WFE for short. A WFE is recursively defined as follows:

1. A variable is a WFE.

2. A constant is a WFE.

3. If E_1 and E_2 are WFE's, then so are $E_1 + E_2$ and $E_1 - E_2$.

4. If E_1 is either a positive constant or a variable, and E_2 is either a positive constant or a variable, then $E_1 \times E_2$ and $E_1 \div E_2$ are WFE's.

Together with the assumption that the variables that occur in $E_1 \times E_2$ or $E_1 \div E_2$ have positive norm values and actual values, these restrictions guarantee that E_1 and E_2 have positive values, and consequently that the definitions given for the $\times$ and $\div$ operators in table 4.14 are correct. A second restriction is that a variable is not allowed to occur more than once on the right-hand side of one and the same equation. The reason for this restriction will become apparent shortly. This restriction will not be a problem in practice, since one can always rewrite the equations in such a way that multiple occurrences are removed, by creating "intermediary" or auxiliary variables. The translation of equations to qualitative relations takes place as follows. Suppose the model M contains an equation $Y = f(X_1, \ldots, X_n)$ that satisfies the above restrictions, i.e., $f(X_1, \ldots, X_n)$ is a WFE and all variables occur at most one time. The influence, pos or neg, of a variable X_i on Y is determined by replacing X_i by the value *high* and all other variables and constants in $f(X_1 \ldots, X_n)$ by the value *normal*. Now there are two possibilities:

1. $f(\partial X_1, \ldots, \partial X_n)$ evaluates to *high*; in this case the influence of X_i on Y is positive.

2. $f(\partial X_1, \ldots, \partial X_n)$ evaluates to *low*; in this case the influence of X_i on Y is negative,

where the expression $f(\partial X_1, \ldots, \partial X_n)$ is evaluated according to the definitions of the operators $+, -$, $\div$ and $\times$ given in table 4.14. Notice that if we had allowed multiple occurrences of a variable, the expression $f(\partial X_1, \ldots, \partial X_n)$ might evaluate to "?". In that case we would not be able to infer a unique direction of influence for the variable concerned. We summarize the translation process by saying that M' is the *qualitative* *linearization* of M, in other words, M' contains a linearized qualitative relation for every equation, linear or non-linear, that occurrs in M.

Note that the seven equations in the interfirm comparison model all have right-hand sides that are WFE's. Translation of these seven equations yields the following qualitative relations:

1. ROA $\leftarrow$ {(PM, pos), (TA, pos)}
2. PM $\leftarrow$ {(AC, neg), (PC, neg), (RDC, neg), (DMC, neg)}
3. TA $\leftarrow$ {(OA, neg)}
4. OA $\leftarrow$ {(CA, pos), (FA, pos)}
5. PC $\leftarrow$ {(MC, pos), (WLC, pos), (WS, pos), (OPC, pos)}
6. CA $\leftarrow$ {(FGS, pos), (DEB, pos), (MS, pos), (WIP, pos)}
7. FA $\leftarrow$ {(LB, pos), (PAM, pos), (OFA, pos)}

The qualitative state of a set of variables is denoted by $\{\partial X_1 = q_1, \ldots, \partial X_n = q_n\}$. The qualitative state QS of the set of all variables that appear in business model M' is called admissible with respect to M' iff for all $Y \leftarrow \{(X_1, s_1), \ldots, (X_m, s_m)\} \in M', \partial Y^{QS} \approx \partial X_1^{QS} \otimes s_1 + \ldots + \partial X_m^{QS} \otimes s_m$. In this expression ∂X_i^{QS} denotes the qualitative value of X_i in qualitative state QS, $\approx$ denotes weak equality as defined in table 4.15, and $\otimes$ is defined in table 4.6. Thus a qualitative state that is admissible with respect to M' is a value assignment to all variables in M' such that all relations in M' are satisfied under weak equality.

4.3.2 Qualitative explanation and diagnosis

In this section we define the notions of *explanation* and *diagnosis* for situations with incomplete information. First, we define the notion of elementary explanation.

Definition 4.11 (Elementary explanation) $\partial X_i = q_i$ *is an elementary explanation for* $\partial Y = q_y$ *iff* $Y \leftarrow \{(X_1, s_1), \ldots, (X_i, s_i), \ldots, (X_m, s_m)\} \in M'$ *and* $\partial X_i \otimes s_i = \partial Y$.

For example, ∂PM = low is an elementary explanation for ∂ROA = low because:

1. ROA $\leftarrow$ {(PM, pos), (TA, pos)} $\in M'$, and

2. low $\otimes$ pos = low.

In words: profit margin (PM) being low is an elementary explanation for return on assets (ROA) being low, because PM occurs in the right-hand side of a relation for ROA and when PM is low, and all other right-hand side variables are normal, then ROA is low. The notion of elementary explanation given in definition 4.11 coincides with the notion of contributing cause for qualitative relations in definition 4.6. Counteracting causes do not appear in the definition of an elementary explanation. Since we are dealing with a situation of incomplete information, explanations will in many cases be hypotheses. If it would also be required to hypothesize possible counteracting causes, this could lead to a very large number of alternative explanations. Furthermore, it makes sense to mention counteracting causes only if they are actually *known* to have occurred, and not as a hypothesis.

Elementary explanations can be chained into explanatory sequences.

Definition 4.12 (Explanatory sequence) *An explanatory sequence for $\partial X_n = q_n$ is a sequence $\partial X_1 = q_1 \rightarrow, \ldots, \partial X_{n-1} = q_{n-1} \rightarrow \partial X_n = q_n$ such that each member of the sequence explains its successor. A* maximal *explanatory sequence for $\partial X_n = q_n$ is an explanatory sequence that starts with a variable that cannot be explained in M' (does not appear on the left-hand side of a model relation).*

For example, ∂EX = high $\rightarrow$ ∂DMC = high $\rightarrow$ ∂PM = low $\rightarrow$ ∂ROA = low is an explanatory sequence for ∂ROA = low. Furthermore it is a *maximal* explanatory sequence since the percentage of export sales (EX) cannot be explained in M'.

Before we define the concept of diagnosis, we should discuss problem identification for diagnosis with incomplete information. It is identical to problem identification for diagnosis with complete information, with the additional requirement that all symptoms have been obtained by comparison with the same reference object. This means that all symptom variables have been obtained by application of the same norm type, e.g. industry average or budget. This requirement is essential for the determination of the mutual consistency of explanatory hypotheses for these symptoms. Thus a set of symptoms is denoted by S^n, where n denotes the normtype concerned. Similarly, we denote the set of known qualitative values with respect to a particular norm type by I^n. To every explanatory sequence E: $\partial X_1 = q_1 \rightarrow \ldots \rightarrow \partial X_n = q_n$, corresponds a partial qualitative state $E^* = \{\partial X_1 = q_1, \ldots, \partial X_n = q_n\}$.

A diagnostic problem is characterised by the following elements:

- M': the set of relations in the transformed business model
- S^n: the set of symptoms to be explained
- I^n: available information about qualitative values of variables in M'

Note that S^n is always a subset of I^n.

Definition 4.13 (Candidate diagnosis) *A candidate diagnosis CD for the triple $< M, I^n, S^n >$ is a set of maximal explanatory sequences such that:*

1. *CD contains at least one maximal explanatory sequence for every symptom in S^n, and*
2. *There exists a qualitative state QS that is admissible wrt M' such that $QS \supseteq \bigcup\{E^* | E \in CD\} \cup I^n$*

The first requirement is fairly straightforward: a diagnosis should explain all symptoms observed. The second requirement demands that the value assignments resulting from the explanatory sequences in CD and known information I^n, can be extended to a qualitative state QS that is admissible with respect to M'.

In general there may be many candidate diagnoses for one and the same diagnostic problem $< M', I^n, S^n >$. We require some selection criteria in order to obtain the most plausible candidates. Normally in model based diagnosis one prefers among all possible explanations those that fulfil some minimality criteria ([Rei87, RNW84]). Reiter calls this *the principle of parsimony*: a diagnosis is a conjecture that some minimal set of components are faulty. This principle follows from the observation that the system description represents the way in which the system *normally* behaves.

However, such a principle of parsimony is less plausible in the business domain. One reason is that the norms, such as historical data and industry averages, are usually only of a heuristic nature. Therefore these norms are less strict than those in the domain of technical devices and the medical domain. Furthermore it should be noted that although reasoning only makes use of qualitative values, the underlying variables are quantitative. Therefore we should consider the situation where several influences jointly produce a particular effect.

For the reasons mentioned above we have decided not to use a principle of parsimony to select the preferred diagnoses from the candidates. Instead we select from the candidates those diagnoses that are *maximal*.

Definition 4.14 (Diagnosis) *A candidate diagnosis CD for $< M', I^n, S^n >$ is maximal iff there is no candidate diagnosis CD' such that $CD \subset CD'$. Such a maximal candidate is called a diagnosis for $< M', I^n, S^n >$.*

Intuitively, a diagnosis contains as many maximal explanatory sequences for the symptoms observed as are mutually consistent. Every candidate diagnosis is a subset of a diagnosis, so a diagnosis conjectures as many explanations as possible for the symptoms observed.

We consider an example of a diagnostic problem using the IFC' model which is the transformed version of the IFC model in section 4.2.3. The set I^{avg} of known qualitative values is presented in table 4.16. Furthermore it is given that problem identification has found that return on assets is below the norm, i.e. $S^{avg} = \{\partial \mathrm{ROA} = \mathrm{low}\}$. We are interested in finding the diagnoses for <IFC', $I^{avg}, \{\partial \mathrm{ROA} = \mathrm{low}\}$>, with I^{avg} as given in table 4.16.

variable	actual	norm (=avg)	∂
return on assets (ROA)	0.088	0.137	low
profit margin (PM)	0.087	0.12	low
total assets turnover (TA)	1.01	1.14	low
fixed assets turnover (FA)	0.4	0.4	normal
materials stock (MS)	0.08	0.1	low
administrative costs (AC)	0.07	0.07	normal
production cost (PC)	0.7	0.75	low

Table 4.16: Qualitative values for diagnostic problem

First we will give all maximal explanatory sequences for ∂ROA = low that are consistent with I^{avg}.

E_1: ∂SPP = low $\rightarrow$ ∂WIP = high $\rightarrow$ ∂CA = high $\rightarrow$ ∂OA = high $\rightarrow$ ∂TA = low $\rightarrow$ ∂ROA = low,

E_2: ∂SPP = high $\rightarrow$ ∂DMC = high $\rightarrow$ ∂PM = low $\rightarrow$ ∂ROA = low,

E_3: ∂SPP = high $\rightarrow$ ∂FGS = high $\rightarrow$ ∂CA = high $\rightarrow$ ∂OA = high $\rightarrow$ ∂TA = low $\rightarrow$ ∂ROA = low,

E_4: ∂EX = high $\rightarrow$ ∂DEB = high $\rightarrow$ ∂CA = high $\rightarrow$ ∂OA = high $\rightarrow$ ∂TA = low $\rightarrow$ ∂ROA = low,

E_5: ∂RDC = high $\rightarrow$ ∂PM = low $\rightarrow$ ∂ROA = low,

E_6: ∂EX = high $\rightarrow$ ∂DMC = high $\rightarrow$ ∂PM = low $\rightarrow$ ∂ROA = low

Each individual explanatory sequence is a *parsimonious* explanation for ∂ROA = low. We are interested in *maximal* sets of mutually consistent explanatory sequences. There are two such sets:

$$D_1 = \{E_1, E_4, E_5, E_6\}$$
$$D_2 = \{E_2, E_3, E_4, E_5, E_6\}$$

The reader will note that the root of the difference between D_1 and D_2 is due to the sales and production policy (SPP). D_1 conjectures that SPP is below the norm, which leads to relatively high work in progress (WIP) and relatively low finished goods stock

(FGS). Thus D_1 explains the low return on assets through high work in progress, which (via current assets) leads to high operating assets. This in turn explains low turnover of assets and consequently low return on assets.

D_2 conjectures that SPP is above the norm. In this case low return on assets is explained by high finished goods stock and high distribution and marketing costs. High distribution and marketing costs leads to low profit margin and therefore low return on assets. High finished goods stock leads to low return on assets through the same causal chain as work in progress. To determine which explanatory sequences actually hold true, further information acquisition will be necessary. One can, however, be certain that the actual diagnosis is a subset of either D_1 or D_2.

4.4 Conclusions

In this chapter, we discussed a formal framework for explanation and diagnosis with *complete* and *incomplete* information. For explanation with complete information, the canonical format of aleatory explanations $< a, F, r >$ because Φ, despite Ψ is adapted to the requirements of the business domain. The sets of contributing and counteracting causes are reduced to "parsimonious" sets, in order to avoid the inclusion of insignificant causes. For the determination of contributing and counteracting causes, we developed the "inf-measure" which embodies a kind of *ceteris paribus* reasoning. In the presence of non-additive functions, the method described may not yield valid results when explanations are non-conjunctive. Therefore some additional constraints on the actual and norm values that are used in an explanation, were necessary in order to enforce the conjunctiveness of explanations. Comparison of the explanations generated by our method with textbook explanations shows that in most cases they are identical. Our method does not have the severe limitation of Courtney et al.'s system, which can only handle linear models. Furthermore, we improve upon Kosy and Wise's method, by eliminating the shortcomings that we pointed out in section 4.2.1.

For the purpose of explanation with incomplete information, the canonical format is effectively reduced to $< a, F, r >$ because Φ, since only contributing causes are hypothesized. The inclusion of counteracting causes in an explanation would only make sense if they were actually *known* to have occurred. Equations in the business model are translated to qualitative relations. Underlying this translation process are a number of assumptions that effectively convert non-additive functions into additive qualitative relations. These assumptions obviously limit the applicability of the method, but they do not seem to be unduly restrictive for the business domain. In our opinion, the "principle of parsimony", which pervades all AI theories of diagnosis, is not applicable to diagnosis of business performance. This reflects the different status of "normal behaviour" in this domain, since normal behaviour does not coincide with the statistically most likely behaviour. One might object that it may not be very likely that actual and norm behaviour are exactly equal, but that it is very likely that actual behaviour is within certain bounds of normal behaviour. The problem with this reasoning is that we should not exclude the possibility of several causes working together to produce a particular effect.

In this chapter we described a formal framework for diagnosis in the business domain. The next logical step is to implement the framework in a computer program. In the next chapter we shall describe the implementation in a logic programming language of diagnosis with complete information and diagnosis with incomplete information.

Chapter 5

Implementation of the formal framework

5.1 Introduction

In this chapter we discuss the implementation of two diagnostic programs. These two programs correspond to the diagnostic problems distinguished in chapter 4: diagnosis with *complete* information and with *incomplete* information respectively.

These two situations reflect basically different problems: the first one deals with explanation as an information selection process, the second one deals with explanation as hypothesis formation. This difference has consequences for the requirements that should be met by a suitable implementation language. In section 5.2 we briefly discuss logic programming and constraint logic programming, and motivate our choice of implementation language for the two programs. In section 5.3 we describe the implementation of the formal framework for diagnosis with complete information. In section 5.4 we describe the implementation of a program for diagnosis with incomplete information. Finally, in section 5.5 we draw a number of conclusions concerning the implementation of the formal framework.

5.2 Logic Programming and Constraint Logic Programming

Logic programming began in the early 1970s as a direct outgrowth of earlier work in automatic theorem proving and artificial intelligence. The fundamental idea underlying logic programming, acredited to Kowalski ([Kow74]) and Colmerauer ([CKRP73]), is that *logic can be used as a programming language*. One of the most important practical outcomes of the research on logic programming is the language PROLOG (PROgramming in LOGic), which is based on the Horn clause subset of logic ([Llo84]). Since the beginning of the 1980s, PROLOG has become one of the main languages for artificial intelligence programming. The properties that make PROLOG a suitable language for AI programming are its declarativeness and its non-determinism, which precludes the need for programming a search procedure. Unfortunately, PROLOG can also be very inefficient when confronted with a natural formulation of constrained search problems ([VH91]). This inefficiency is mainly caused by the passive use of constraints, which only *tests* potential values, instead of pruning the search space in an active manner. As a consequence, formulation of a constrained search problem in PROLOG often leads to "generate and test" or "standard backtracking" behaviour ([VH91]).

Constraint Logic Programming (CLP) languages replace *unification*, which is the basic operation of PROLOG, by constraint solving in some computation domain, for example the domain of real numbers. Unification can be regarded as a simple form of constraint solving, namely solving equations among first-order terms, and thus CLP can be viewed as a generalisation of Logic Programming. By providing efficient constraint-solving methods (from algebra, artificial intelligence, operations research, and logic), CLP languages support, in a declarative way, computational paradigms such as partial enumeration, constraint satisfaction, and branch and bound. The resulting languages combine the advantages of logic programming, namely declarative semantics, relational form and nondeterminism, with the efficiency of special purpose constraint solvers. Various CLP languages such as CHIP ([DHS+88]), CLP($\mathcal{R}$) ([JMSY92]), and PROLOG-III ([Col90]) have been implemented in recent years.

At first sight it may seem that PROLOG is *not* such a suitable language for the implementation of the formal framework for diagnosis with complete information, since this problem requires some amount of numerical computation in order to determine the influence of each variable, and consequently to determine a diagnosis. The power of a language like PROLOG for this problem primarily lies in the fact that the business model can be viewed both as a set of equations, which can be evaluated numerically, i.e. the business model is a computing model, and as a set of PROLOG terms, which can be inspected and manipulated symbolically. Thus, given the proper representation of the business model, it can be viewed both in a declarative way, as a "knowledge structure", and in a procedural way, as arithmetic expressions that can be evaluated. In typed languages such as C or Pascal this dual functionality would yield considerably more problems. This point will become clearer in the next section.

For the implementation of the formal framework for diagnosis with incomplete information, PROLOG's inefficiency in solving constrained search problems becomes a severe disadvantage. The formation of consistent explanatory hypotheses can be viewed as a constrained search problem in a very natural way. Therefore we chose to implement the system in a constraint logic programming language. Since explanation takes place in a quantity space, that contains a finite number of qualitative values, we selected a language that is able to deal with constraints over finite domains. Therefore the constraint logic programming language CHIP ([DHS+88]) was chosen. CHIP is a CLP language developed at ECRC in Munich, and has been designed to tackle real-world constrained search problems. In CHIP, three new computation domains are introduced, namely finite domains, booleans and linear rational terms. For each of these domains CHIP uses specialized constraint solving techniques: consistency techniques for finite domains, equation solving in boolean algebra for booleans, and a symbolic simplex-like algorithm for rationals. For details, the reader is referred to ([DHS+88, VH91, VH89]).

In the next two sections we discuss the implementations of diagnosis with complete information and diagnosis with incomplete information respectively.

5.3 Prolog implementation of diagnosis with complete information

In this section we discuss the implementation of the formal framework for diagnosis with complete information in the logic programming language PROLOG. Figure 5.1 depicts the architecture of the program.

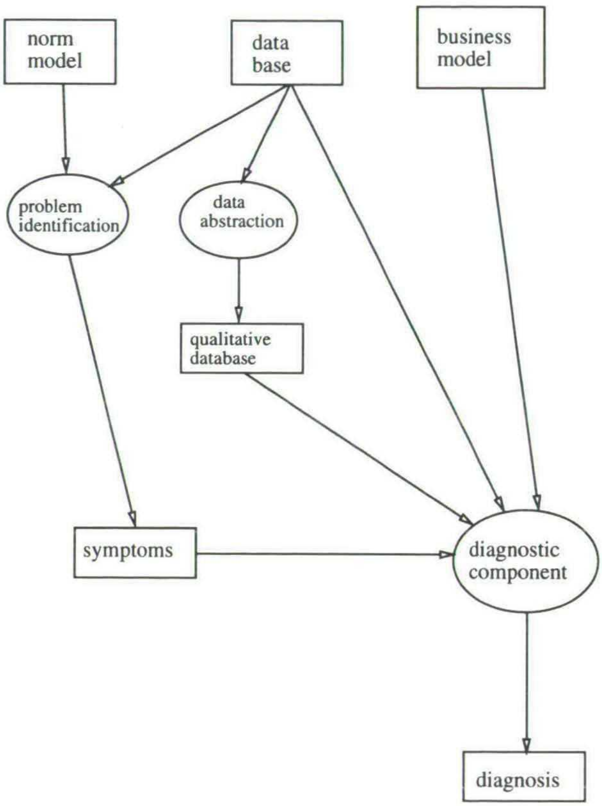

Figure 5.1: Architecture of program for diagnosis with complete information

We discuss the implementation of different elements of the program, and conclude with an example of the program's analysis of the interfirm comparison case that was presented in section 4.2.3.

5.3.1 Representation of the business model

As we have already stressed in the introduction, the business model is represented in such a way that it can be viewed in both a declarative way and a procedural way. This effectively allows the business model to be a readable knowledge base. First we illustrate the representation of a business model by showing the program representation of the interfirm comparison model of section 4.2.3.

```
model([roa = pm * ta,
       pm = 1 - (ac + pc + rdc + dmc),
       ta = 1 / oa,
       oa = ca + fa,
       pc = mc + wlc +  ws + opc,
       ca = ms + wip + fgs + deb,
       fa = lb + pam + ofa,
       rel(wip,[inf(spp,neg)]),
       rel(mc,[inf(mu,pos),inf(bc,pos)]),
       rel(dmc,[inf(spp,pos),inf(ex,pos)]),
       rel(fgs,[inf(spp,pos)]),
       rel(deb,[inf(ex,pos)])]).
```

As this example shows, the predicate model/1[1] contains as its single argument a list of equations and qualitative relations, together forming the business model. All variables appearing in the business model are represented as, and uniquely identified by, PROLOG atoms. An equation has the following general structure:

$< atom > = < term >$,

where $< term >$ is a PROLOG term consisting of atoms and the operators +,−,/, and *. The representation of qualitative relations is also straightforward: each relation is represented as

rel($< lhs >$,$< loi >$),

where $< lhs >$ denotes the left-hand side variable of the relation, and $< loi >$ denotes a list containing all right-hand side variables together with the sign, pos or neg, of their respective influences.

[1]The number following the name of a predicate indicates its arity.

5.3.2 Representation of the data base

The predicate database/3 is used to represent the data that will be used for diagnosis. The following clause represents the 1990 database for firm 8 from the interfirm comparison case.

```
database(1990,firm8,[
 roa = 0.079,
 pm = 0.089,
 ta = 0.9,
 oa = 1.111,
 pc = 0.774,
 rdc = 0.0,
 dmc = 0.046,
 ac = 0.091,
 mc = 0.347,
 wlc = 0.274,
 opc = 0.153,
 ws = 0.0,
 ca = 0.582,
 fa = 0.529,
 ms = 0.101,
 wip = 0.225,
 fgs = 0.053,
 deb = 0.203,
 lb = 0.277,
 pam = 0.246,
 ofa = 0.006,
 ex = 0.12,
 spp = 1,
 bc = 0.37,
 mu = 2]).
```

The first argument of database/3 represents the time period that the data apply to. The second argument is used either for the company name or for some other identifying label, e.g. budget or industry average. The third argument is a list containing the data in the format $< variable >=< value >$.

5.3.3 Problem identification

In section 3.2 we specified the structure of a unit in the norm model. An example of the representation of a norm unit for the interfirm comparison model in the program is:

```
norm-unit(variable(roa),
          normtype(ind_avg),
          function(rd),
          low_bound(-0.05),
          up_bound(1000)).
```

This **norm-unit** specifies that (1) the normtype for **roa** is its industry average, which is indicated by the value **ind_avg** in the **normtype** slot, and (2) the function to be used to compute the deviation between actual and norm values is their relative difference **rd**. The value -0.05 in the **lowbound** slot indicates that a relative difference of less than -0.05 is considered to be a significant deviation for **roa**. The value 1000 in the **upbound** slot practically implies that a value of **roa** that is above the norm, is never considered as a symptom.

The predicate for identifying symptoms implements problem identification as described in section 4.2.2:

```
symptom(Year,Company,s(V,N)):-
  norm-unit(variable(V),
            normtype(N),
            function(F),
            low_bound(L),
            up_bound(U)),
  values(Year,Company,V,N,Actual,Norm),
  deviation(F,Actual,Norm,Dev),
  compare(Dev,L,U,s(V,N)).

compare(Dev,L,U,s(V,N)):-
  Dev < L.

compare(Dev,L,U,s(V,N)):-
  Dev > U.
```

The **values** predicate returns the actual value and norm value of the variable specified in the norm unit. The **deviation** predicate returns the result of applying the function specified in the function-slot of the norm unit, to the actual value and norm value of variable **V**. Finally, the **compare** predicate compares the function result with the upperbound and lowerbound specified in the `norm-unit` and discovers a symptom when appropriate.

The predicate **symptom/3** should be called as follows:

```
symptom(+Year,+Company,-Symptom)
```

where **+X** denotes an input argument and **-X** denotes an output argument. An input argument is an argument that must be instantiated when the predicate is called. An output argument is an argument that must *not* be instantiated when the predicate is called. In the example, when symptom/3 is called, the **Year** and **Company** must be specified; the third argument must be an uninstantiated variable, that returns a symptom when appropriate. By backtracking over all norm units in the norm model, all symptoms for the year and company specified are eventually found.

5.3.4 Explanation generation

The predicate for the generation of explanations is at the heart of the diagnostic program. For diagnostic purposes this predicate is called in order to explain significant differences between norms and actual values, discovered by the problem identification module. The predicate can also be used in a more interactive manner to explain any difference the user may be interested in. The predicate explains/6 is called as follows:

```
explains(+Lhs,+T,+c(Y1,C1),+c(Y2,C2),-Cbs,-Cas).
```

The first argument **Lhs** denotes the model variable that the explanation is concerned with. The second argument denotes the fraction T that should be explained. The third and fourth arguments identify the two "objects" to be compared by stating the respective years, **Y1** and **Y2**, and company names or other identifying labels, **C1** and **C2**. All these arguments should be instantiated when explains/6 is called. The fifth and sixth arguments denote the lists of parsimonious contributing and counteracting causes respectively. They must *not* be instantiated when explains/6 is called.

The set of parsimonious contributing and counteracting causes is determined as follows. In case a quantitative equation is involved, the program first determines the sets of contributing and counteracting causes. Recall from section 4.2.2 that the sets of *parsimonious* contributing and counteracting causes were defined as the smallest subsets of their complete counterparts that explain more than some predefined fraction T of the complete set. One way to compute these sets is simply to test all subsets in ascending order of cardinality until one is found that exceeds T. This method is guaranteed to work and is fairly inefficient since a large number of subsets may have to be tested. In the current version of the program this method has been replaced by a more efficient one; however, in some pathological cases it may not yield the smallest subset. The sets of parsimonious contributing and counteracting causes are constructed by iteratively adding the element from the complete set with the largest inf-value. This iteration continues until the fraction T has been reached. In case a qualitative relation supports the explanation, the 5th and 6th argument return the *complete* set of contributing and counteracting causes respectively. The causes are determined in a qualitative manner, according to the description given in section 4.2.2, definitions 4.6 and 4.7.

The predicate `max_explanation/5` returns the maximal explanation, according to definition 4.8, for a symptom that has been discovered by the problem identification predicate. Basically it is a recursive version of the explains/6 predicate.

```
max_explanation([Lhs|R1],T,c(Y1,C1),c(Y2,C2),[Lhs+Cbs,Lhs-Cas|R2]) :-
   explains(Lhs,T,c(Y1,C1),c(Y2,C2),Cbs,Cas),
   append(Cbs,R1,R3), % depth-first traversal
   max_explanation(R3,T,c(Y1,C1),c(Y2,C2),R2).

max_explanation([],_,_,_,[]). % stop condition
```

Starting with a symptom, it recursively applies explains/6 to its contributing causes. Since the contributing causes of a variable are added to the beginning of the list, we obtain a depth-first generation of the maximal explanation. This tree of contributing and counteracting causes is returned by the fifth argument, where `Lhs+Cbs` denotes the set of contributing causes and `Lhs-Cas` denotes the set of counteracting causes of `Lhs`. The recursion stops when the list of contributing causes, i.e. the first argument, is empty.

That case is handled by the second `max_explanation` clause.

5.3.5 Example of program use

In this section we illustrate the use of the program by showing an example diagnosis for firm 9 from the interfirm comparison case. The top-level predicate is analyse/4, which should be called as follows:

```
analyse(+Year, +Company, -Symptoms, -Diagnosis).
```

We illustrate its use by the following query. This query corresponds to the diagnosis for firm 9 in the interfirm comparison case, which we discussed in section 4.2.3.

```
| ?- analyse(1990,firm9,S,D).

I have discovered the following symptoms:
  roa at firm9 is lower than roa at ind_avg.

roa at firm9 is lower than roa at ind_avg.
The major contributing causes are:
  pm at firm9 is lower than pm at ind_avg.
  ta at firm9 is lower than ta at ind_avg.
There are no counteracting causes.

pm at firm9 is lower than pm at ind_avg.
The major contributing causes are:
  pc at firm9 is higher than pc at ind_avg.
The major counteracting causes are:
  dmc at firm9 is lower than dmc at ind_avg.
  rdc at firm9 is lower than rdc at ind_avg.

pc at firm9 is higher than pc at ind_avg.
The major contributing causes are:
  wlc at firm9 is higher than wlc at ind_avg.
  opc at firm9 is higher than opc at ind_avg.
  mc at firm9 is higher than mc at ind_avg.
The major counteracting causes are:
  ws at firm9 is lower than ws at ind_avg.

There is no equation for wlc in the model.
There is no equation for opc in the model.

mc at firm9 is higher than mc at ind_avg.
```

```
The major contributing causes are:
  bc at firm9 is higher than bc at ind_avg.
The major counteracting causes are:
  mu at firm9 is lower than mu at ind_avg.

There is no equation for bc in the model.

ta at firm9 is lower than ta at ind_avg.
The major contributing causes are:
  oa at firm9 is higher than oa at ind_avg.
There are no counteracting causes.

oa at firm9 is higher than oa at ind_avg.
The major contributing causes are:
  fa at firm9 is higher than fa at ind_avg.
  ca at firm9 is higher than ca at ind_avg.
There are no counteracting causes.

fa at firm9 is higher than fa at ind_avg.
The major contributing causes are:
  pam at firm9 is higher than pam at ind_avg.
  lb at firm9 is higher than lb at ind_avg.
There are no counteracting causes.

There is no equation for pam in the model.
There is no equation for lb in the model.

ca at firm9 is higher than ca at ind_avg.
The major contributing causes are:
  wip at firm9 is higher than wip at ind_avg.
The major counteracting causes are:
  ms at firm9 is lower than ms at ind_avg.
  fgs at firm9 is lower than fgs at ind_avg.

wip at firm9 is higher than wip at ind_avg.
The major contributing causes are:
  spp at firm9 is lower than spp at ind_avg.
There are no counteracting causes.

There is no equation for spp in the model.

S=[s(roa,ind_avg)]

D =[[roa+[pm,ta],roa-[],pm+[pc],pm-[dmc,rdc],pc+[wlc,opc,mc],pc-[ws],wlc+[],
wlc-[],opc+[],opc-[],mc+[bc],mc-[mu],bc+[],bc-[],ta+[oa],ta-[],oa+[fa,ca],
```

```
oa-[],fa+[pam,lb],fa-[],pam+[],pam-[],lb+[],lb-[],ca+[wip],ca-[ms,fgs],
wip+[spp],wip-[],spp+[],spp-[]]] ;
```

Apart from returning the instantiations of `S` and `D` for which the query succeeds, the program generates text output. Firstly, the text output lists all symptoms that have been discovered, in this case only one. Secondly, it generates text corresponding to the maximal explanation of each symptom. The textual explanation traverses the maximal explanation in a depth-first manner.

5.4 CHIP implementation of diagnosis with incomplete information

In the following sections, a description is given of the different elements of the CHIP program for diagnosis of business performance with incomplete information. The architecture of this program is depicted in figure 5.2.

5.4.1 Quantitative constraint solving

Quantitative constraint solving involves the equations of the business model. It uses the available quantitative data, to infer as many values as possible. For this purpose, the program uses one of the specialised computation domains of CHIP, namely linear rational terms. CHIP uses a symbolic simplex-like algorithm to solve linear constraints over rational terms. The program delays a non-linear constraint until enough variables have been instantiated to make it linear. This is achieved by using the built-in delay mechanism of CHIP. Compared with spreadsheet programs, which are often used for this type of business calculations, CHIP has the advantage that it is not required to specify in advance which variables are *input* to the system and which variables have to be computed. This corresponds to what Fordyce ([For87]) calls *reversibility* of the model. Since the constraints do not have to be specified in any particular order, the declarative reading of the program is preserved. After inferring as many quantitative values as possible, the results are passed to (1) the problem detection module in order to identify symptoms that need to be explained and (2) the data-abstraction module in order to

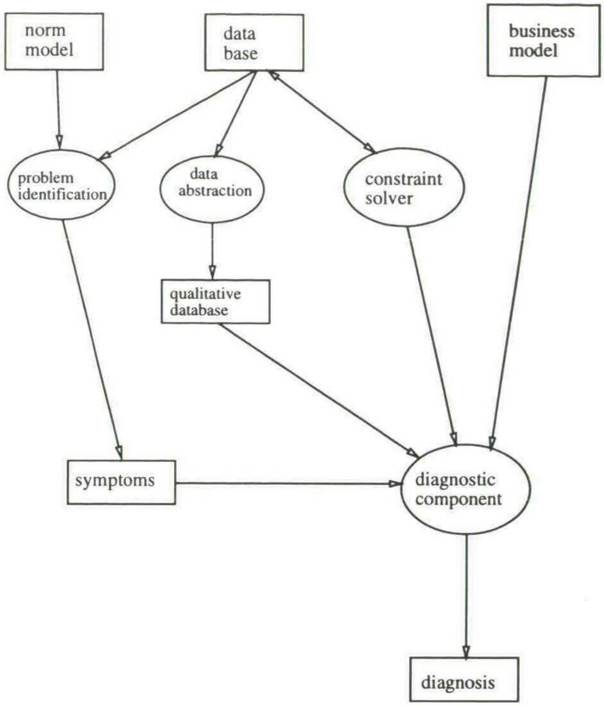

Figure 5.2: Architecture of program for diagnosis with incomplete information

translate the data into qualitative terms. We will not discuss problem identification here, since it is virtually identical to problem identification for the complete information case, which we have already described in the previous section. We shall discuss qualitative data abstraction next.

5.4.2 Qualitative data abstraction

After quantitative constraint solving is finished, the program makes a qualitative abstraction of the data. Data abstraction is required for diagnosis, which is performed in a purely qualitative manner. This process is identical to the data abstraction process specified in table 4.3, except for the fact that the qualitative values are represented in the program by integer values as shown in table 5.1. The result of qualitative data abstraction is called the qualitative data base. The qualitative data base is represented as

Qualitative value	Program representation
high	2
normal	1
low	0

Table 5.1: Program representation of qualitative values

a list of terms of the form `Var = Qvalue`, where `Var` denotes a variable and `Qvalue` its qualitative value.

5.4.3 Representation of constraints for diagnosis

For the purpose of diagnosis with incomplete information, all relations are interpreted qualitatively. The qualitative relations have to be entered into the program in this format, whereas for the quantitative relations a translation from equational form is provided. This translation is performed according to the rules presented in section 4.3.1. As an example, we show the program representation of the qualitative relations, corresponding to the interfirm comparison model:

```
model([rel(Roa,[inf(Pm,pos),inf(Ta,pos)]),
       rel(Pm,[inf(Ac,neg),inf(Pc,neg),inf(Rdc,neg),inf(Dmc,neg)]),
       rel(Ta,[inf(Oa,neg)]),
       rel(Oa,[inf(Ca,pos),inf(Fa,pos)]),
       rel(Pc,[inf(Mc,pos),inf(Wlc,pos),inf(Ws,pos),inf(Opc,pos)]),
       rel(Ca,[inf(Fgs,pos),inf(Deb,pos),inf(Ms,pos),inf(Wip,pos)]),
       rel(Fa,[inf(Lb,pos),inf(Pam,pos),inf(Ofa,pos)]),
       rel(Wip,[inf(Spp,neg)]),
       rel(Mc,[inf(Mu,pos),inf(Bc,pos)]),
       rel(Dmc,[inf(Spp,pos),inf(Ex,pos)]),
       rel(Fgs,[inf(Spp,pos)]),
       rel(Deb,[inf(Ex,pos)])]).
```

5.4.4 Diagnosis

In the diagnostic component, potential explanations for the symptoms are constructed. First it is shown how explanatory sequences for a symptom are generated and how candidate diagnoses are constructed from them. Thereafter we discuss the use of qualitative constraint solving to determine the acceptability of a candidate diagnosis.

The explains/4 predicate generates maximal explanatory sequences for the symptoms discovered:

```
% explains(+Symptom,-ExplanatorySequence,+Constrainst,+QValues)

explains(Var1 = QValue1,[Var2 = QValue2|Rest],Constraints,QValues):-
 member(rel(Var1,InfList),Constraints),
 member(inf(Var2,Inf),InfList),
 result(QValue2,Inf,QValue1),
 not (inconsistent(Var2 = QValue2,QValues)),
 explains(Var2 = Value2,Rest,Constraints,QValues).

explains(Var = QValue,[],Constraints,QValues):-
 not (member(rel(Var,InfList),Constraints)).

result(Qvalue,pos,Qvalue).

result(Qvalue1,neg,Qvalue2):-
 inv(Qvalue1,Qvalue2).

inv(0,2).
inv(2,0).

inconsistent(Var = QValue1,QValues):-
 member(Var = QValue2,QValues),
 QValue2 \== QValue1.
```

Here the argument `Constraints` is a list containing all constraints in the business model and `QValues` is a list containing all variable-qualitative value pairs known to the system, i.e. the qualitative data base. In order to generate a maximal explanatory sequence for a symptom, the `explains` predicate retrieves the appropriate constraint from the business model. The `inconsistent` predicate determines if an explanation is in direct contradiction with the values in the qualitative data base. If this is the case, the explanatory sequence in which the explanation occurs is not generated. The second `explains` clause is the stopping condition, which succeeds if there is no constraint in the business model with the variable to be explained "on the left-hand side".

The program generates *pseudocandidates* by taking all different combinations of maximal explanatory sequences for the symptoms discovered such that

- every pseudocandidate contains *at least one* maximal explanatory sequence for every symptom;

- every pseudocandidate C is internally consistent: there are no two maximal explanatory sequences $E_1, E_2 \in C$ that assign different values to the same variable.

A *pseudocandidate* is similar to a *candidate diagnosis*, as defined in definition 4.13, except that it has not been checked for global consistency. The total set of pseudocandidates is sorted in descending order of cardinality. The following predicate returns all diagnoses:

```
% max_cand(+PseudoCandidates,-Diagnoses,+NormType,+Vars,+Constraints)

max_cand([],[],_,_,_).

max_cand([C|Cs1],[C|Ds],NormType,Vars,Constraints):-
  diagnose(C,NormType,Vars,Constraints),!,
  removesub(C,Cs1,Cs2),
  max_cand(Cs2,Ds,NormType,Vars,Constraints).

max_cand([C|Cs],Ds,NormType,Vars,Constraints):-
  max_cand(Cs,Ds,NormType,Vars,Constraints).
```

The first argument of `max_cand` is the ordered list of pseudo candidates. The second argument is the list of diagnoses that is returned by the predicate. The first clause succeeds when the empty list of pseudocandidates has been reached. The second clause succeeds if the first element of the pseudocandidate list is globally consistent, which is determined by the `diagnose` predicate. The `removesub` predicate succeeds if list `Cs2` is equal to `Cs1` with all subsets of `C` removed. Consequently, only the *maximal* candidates are eventually returned. The third clause succeeds if `C` is not globally consistent. In that case the pseudocandidate is removed from the head of the list, and the predicate continues with the remaining pseudocandidates.

A pseudocandidate together with the qualitative data base forms a partial value assignment to the variables in the constraints. Qualitative constraint solving is used to determine the global consistency of a pseudocandidate. The task is to find *one* solution to the constraints. For this purpose the model variables are represented as *finite domain* variables, ranging over the domain 0..2. The predicates for constraint solving read as follows:

```
% apply-constraints(+Constraints,?Variables)

apply-constraints([],Vars).
apply-constraints([rel(Y,InfList)|Constraints],Vars):-
 evaluates(Y,InfList,Vars),
 apply-constraints(Constraints,Vars).

evaluates(Y,[inf(X,I)],Vars):-
 otimes(X,I,Y),!.

evaluates(Y,[inf(X,I)|Infs],Vars):-
 Y1::0..2, % representation of Y1 as a finite domain variable
 Y2::0..2,
 otimes(X,I,Y1),
 evaluates(Y2,T,Vars),
 add(Y1,Y2,Y).

otimes(Y,pos,Y).

otimes(Y,neg,Z):-
 Y + Z <=> 2.
```

The first argument of the **apply-constraints** predicate is a list of model constraints. The second argument is a list of **[Atom,Variable]** pairs, e.g. **[advertising,Advertising]**. Because this list is passed along during constraint solving, the variables in this list get instantiated as soon as the same variable in the constraints gets instantiated. Thus the results of constraint solving are eventually stored in this list. The **evaluates** predicate recurses on **InfList** and calls the **otimes** and **add** predicate. The **<=>** symbol appearing in the **otimes** predicate denotes equality over finite domain linear terms. The **otimes** predicate implements the $\otimes$-operator as defined in chapter 4, table 4.6. The **add** predicate is at the heart of all qualitative constraint solving and its efficient implementation therefore has substantial influence on the overall efficiency of the program. It has been implemented using the built-in demon construction of CHIP. Its definition is as follows:

```
?- demon add/3. % declaration of add as a demon predicate

add(A,1,C):- A = C.
add(1,B,C):- B = C.
add(A,B,1):- A + B <=> 2.
add(0,B,2):- B = 2.
add(2,B,0):- B = 0.
add(A,2,0):- A = 0.
add(A,0,2):- A = 2.
add(A,A,C):- A = C.
```

The definition of a predicate as a demon predicate has important consequences for the way it is used in resolving a goal:

- A goal can only be resolved against a demon predicate if it *matches* a head in the definition clauses. Matching is unification without binding any variables in the goal clause.
- When the goal matches one of the heads, it cannot backtrack to match with another one, i.e. the predicate is *deterministic*.
- If a goal does not match any head, it is delayed until more variables in the goal clause are instantiated.

The point of using a demon predicate is to avoid making choices for variables too soon, e.g. when all variables in the goal clause are still uninstantiated. By delaying these goals until more variables are instantiated, the program can avoid costly backtracking. Suppose, for example, that at a specific point in the computation the goal to be resolved is `add(2,B,C)`. This goal does not match any demon head and hence will be delayed until either B or C gets instantiated. Had the add-constraint been represented as a set of PROLOG facts, i.e. the program contains a fact for every tuple for which the constraint holds, then the goal would resolve against one of them. This would create a choice point and consequently the possibility of extensive backtracking.

Finally, the predicate to determine the global consistency of a candidate diagnosis reads as follows:

```
consistent(VA,Vars,Constraints):-
  qunify(VA,Vars,Constraints),
  apply-constraints(Constraints,Vars),
  labeling(Vars),!.
```

The first argument of the `consistent` predicate is a value assigment to the variables in the constraints, resulting from the candidate diagnosis and the qualitative data base. The `qunify` predicate takes care of the proper instantiatiation of variables in the constraints, according to the value assignment `VA`. If after applying all constraints, by `apply-constraints`, there are only delayed add constraints left, the program applies a labeling procedure that chooses a value for a variable from its domain. If this leads to the instantiation of a variable in a delayed constraint, the system checks if it can now apply a demon clause to this constraint. The generation of values by this labeling procedure is non-deterministic and will eventually lead to finding all solutions to the constraints. For the problem at hand it is not required to find all solutions, but only to make sure that there is *at least one*. Therefore the `consistent` predicate does not backtrack after finding the first solution.

5.4.5 Example case of program behaviour

The top-level predicate of the program should be called as follows:

```
?- analyse(+Year,+Company,+NormType,-Symptoms,-Diagnoses).
```

The first argument of this call specifies the year for which the analysis should be made. The second argument denotes the company for which the diagnosis should be performed. The third argument indicates whether the norm values to be used are last year's values, i.e. `NormType = historical`, or the industry average of the variables, i.e. `NormType = ind_avg`. The fourth argument returns the symptoms that have been discovered by the problem identification predicate. The fifth argument finally returns the maximal diagnoses that have been determined by the diagnostic part of the program.

Suppose the program receives the input given in table 5.2, which corresponds to the example given in section 4.3.2. The applicable norm unit is:

variable	firm 10	average
return on assets (ROA)	0.088	0.137
profit margin (PM)	0.087	0.12
total assets turnover (TA)	1.01	1.14
fixed assets turnover (FA)	0.4	0.4
materials stock (MS)	0.08	0.1
administrative costs (AC)	0.07	0.07
production cost (PC)	0.7	0.75

Table 5.2: Program input

```
norm-unit(variable(roa),
          norm(ind_avg),
          function(rd),
          low_bound(-0.05),
          up_bound(1000)).
```

The query:

```
24?- analyse(1990,firm10,ind_avg,S,D).
```

produces the following program output:

```
Financial and operating data for ind_avg in 1990:

Roa     0.137
Pm      0.12
Ta      1.14
Ac      0.07
Pc      0.75
Rdc     unknown
Dmc     unknown
Oa      0.88
Ca      0.44
Fa      0.4
Mc      unknown
Wlc     unknown
Ws      unknown
Opc     unknown
Fgs     unknown
Deb     unknown
Ms      0.1
```

```
Wip    unknown
Lb     unknown
Pam    unknown
Ofa    unknown
Spp    unknown
Mu     unknown
Bc     unknown
Ex     unknown

Financial and operating data for firm10 in 1990

Roa    0.088
Pm     0.087
Ta     1.01
Ac     0.07
Pc     0.7
Rdc    unknown
Dmc    unknown
Oa     0.99
Ca     0.59
Fa     0.4
Mc     unknown
Wlc    unknown
Ws     unknown
Opc    unknown
Fgs    unknown
Deb    unknown
Ms     0.08
Wip    unknown
Lb     unknown
Pam    unknown
Ofa    unknown
Spp    unknown
Mu     unknown
Bc     unknown
Ex     unknown

I have detected the following symptoms:
roa of firm10 is lower than industry average

There are 2 maximal diagnoses.
The first maximal diagnosis is:
roa of firm10 is below industry average is explained by
  ta of firm10 is below industry average, which is explained by
  oa of firm10 is above industry average, which is explained by
```

```
  ca of firm10 is above industry average, which is explained by
  wip of firm10 is above industry average, which is explained by
  spp of firm10 is below industry average.

roa of firm10 is below industry average is explained by
  ta of firm10 is below industry average, which is explained by
  oa of firm10 is above industry average, which is explained by
  ca of firm10 is above industry average, which is explained by
  deb of firm10 is above industry average, which is explained by
  ex of firm10 is above industry average.

roa of firm10 is below industry average is explained by
  pm of firm10 is below industry average, which is explained by
  rdc of firm10 is above industry average.

roa of firm10 is below industry average is explained by
  pm of firm10 is below industry average, which is explained by
  dmc of firm10 is above industry average, which is explained by
  ex of firm10 is above industry average.

The second maximal diagnosis is:
roa of firm10 is below industry average is explained by
  pm of firm10 is below industry average, which is explained by
  dmc of firm10 is above industry average, which is explained by
  spp of firm10 is above industry average.

roa of firm10 is below industry average is explained by
  ta of firm10 is below industry average, which is explained by
  oa of firm10 is above industry average, which is explained by
  ca of firm10 is above industry average, which is explained by
  fgs of firm10 is above industry average, which is explained by
  spp of firm10 is above industry average.

roa of firm10 is below industry average is explained by
  ta of firm10 is below industry average, which is explained by
  oa of firm10 is above industry average, which is explained by
  ca of firm10 is above industry average, which is explained by
  deb of firm10 is above industry average, which is explained by
  ex of firm10 is above industry average.

roa of firm10 is below industry average is explained by
  pm of firm10 is below industry average, which is explained by
  rdc of firm10 is above industry average.

roa of firm10 is below industry average is explained by
```

```
  pm of firm10 is below industry average, which is explained by
  dmc of firm10 is above industry average, which is explained by
  ex of firm10 is above industry average.

End of analysis.

S = [s(roa,0)]
D = [[[v(roa,0),v(ta,0),v(oa,2),v(ca,2),v(wip,2),v(spp,0)],
      [v(roa,0),v(ta,0),v(oa,2),v(ca,2),v(deb,2),v(ex,2)],
      [v(roa,0),v(pm,0),v(rdc,2)],
      [v(roa,0),v(pm,0),v(dmc,2),v(ex,2)]],
     [[v(roa,0),v(pm,0),v(dmc,2),v(spp,2)],
      [v(roa,0),v(ta,0),v(oa,2),v(ca,2),v(fgs,2),v(spp,2)],
      [v(roa,0),v(ta,0),v(oa,2),v(ca,2),v(deb,2),v(ex,2)],
      [v(roa,0),v(pm,0),v(rdc,2)],
      [v(roa,0),v(pm,0),v(dmc,2),v(ex,2)]]] ? ;
no (more) solutions
```

First the program presents the results of the quantitative constraint solving. Then it gives the list of symptoms that were discovered by the program. The problem identification predicate discovers a significant decrease in Return on Assets (roa). Thereafter the maximal diagnoses follow. In this case there are two maximal diagnoses. The text output of the program is rather inflexible and could be improved upon. In the above case, for example, it is possible to summarize the explanatory sequences by making use of their large overlap.

5.5 Conclusions

In this chapter we discussed the implementation of two programs: one for diagnosis with complete information and one for diagnosis with incomplete information. The use of the logic programming language PROLOG and the constraint logic programming language CHIP made it fairly easy to represent the required knowledge in a declarative way. We were able to maintain a strict separation between (1) the easily readable knowledge base, containing the norm model and the business model, and (2) the reasoning part of the program. Furthermore, the built-in search procedure of PROLOG precluded the need for explicit tree programming.

For diagnosis with incomplete information, PROLOG is not very suited, because of its inefficiency in solving constrained search problems. For this reason we implemented the program in the constraint logic programming language CHIP. Because of CHIP's ability to actively use constraints over finite-domain variables, we were able to determine the global consistency of potential diagnoses more efficiently than would have been the case with PROLOG. Furthermore, we used constraints over linear rational terms in order to obtain a "versatile spreadsheet" for the inference of new quantitative values from the program input.

For both programs, the presentation of the output could be improved upon. The current text generation is rather inflexible; for example, it does not take advantage of the presence of a large overlap between explanations to summarize them.

Chapter 6

Summary

The formalisation of diagnostic problem solving is a sub-area of Artificial Intelligence (AI) research that has received considerable attention in recent years. Diagnosis is defined here as finding the best explanation of observed abnormal behaviour of a system under study. Especially "model based", as opposed to "heuristic classification" approaches have been developed and investigated, mainly because of their supposedly superior problem solving and explanation capabilities. The larger part of research into diagnostic problem solving has either implicitly or explicitly been concerned with medical diagnosis or with diagnosis of man-made artifacts such as electronic circuits. This focus has had consequences for knowledge representation formalisms and associated reasoning methods. For example, in the domain of diagnosis of electronic circuits, the system concerned is usually represented as a set of first-order logic sentences, describing the circuit components and the way they are interconnected. In the medical domain, one often encounters causal models, which describe cause-effect relations between "disease states". In accordance with the knowledge available in this domain these causal models are of a qualitative nature.

In this study, however, we have developed a formalisation of diagnostic reasoning in the domain of business and finance, more specifically, diagnosis of business performance. We selected the analysis of the concept *explanation* as the central issue. In chapter 2 we developed a causal view of explanation that can deal with quantitative phenomena, which pervade the domain of business and finance.

With respect to the issue of knowledge representation, our formalisation can be characterised as a *model based* rather than a *heuristic classification* approach. Apart from the advantages ascribed to a model based approach, it also appears to be the most natural approach in the domain of business and finance. On the one hand this is so, because a substantial part of the relevant knowledge deals with relations among financial and operational variables. This knowledge is usually already available in the form of a system of equations, which is used to *compute* values rather than to *explain* them. Such equations often model definitions. On the other hand there is knowledge of "genuine" cause-effect relations that is often of a qualitative nature.

In chapter 4 we developed two different approaches to diagnosis, one based on the assumption of complete information and one based on the assumption of incomplete information. In the second case we were forced to resort to qualitative reasoning, in order to obtain a finite number of explanatory hypotheses. Incomplete information almost seems to be a defining characteristic of diagnostic problems in non-business domains. In our view this is less so in the business domain, where the selection of significant influences in complete information situations is also viewed as diagnosis. In practice the two approaches could be merged into one system, applying qualitative and quantitative reasoning respectively when appropriate. Comparison of the explanations generated by our method with textbook explanations shows that in most cases they are identical. Our method does not have the severe limitation of Courtney et al.'s system, which can only handle linear models. Furthermore, we improve upon Kosy and Wise's method, by eliminating the shortcomings that we pointed out in section 4.2.1.

In chapter 5 we discussed the implementation of two diagnostic reasoners in the constraint logic programming language CHIP and the logic programming language PROLOG respectively. The (constraint) logic programming paradigm was shown to offer the flexibility to view the business model as a declarative knowledge structure, subject to symbolic manipulation, as well as to use it as a computing model.

The diagnostic methods developed were inspired on the one hand by the scarce work already done in the area of automated diagnostic reasoning in business and finance, and on the other hand by a case study that we performed. In appendix A we described this case study, which involves a real-life financial analysis task. We analysed the problem-solving protocols of a stock analyst, who evaluated the performance of a number of German construction companies. Analysis of the protocols revealed a number of properties of problem identification and explanation by financial analysts. These properties have been integrated into our theoretical model, both where it concerns the application of norm values and where it concerns the explanation of significant differences between norm value and actual performance.

Diagnosis of business performance seldomly occurs as an isolated task. In the case of the stock analyst, for example, an assessment of the different companies concerned had to be made, which in turn lead to an ordering of their attractiveness as investment objects. Especially in managerial decision making, diagnosis is but one step in a larger process. The coupling of diagnosis and "therapy", however, did not receive attention in this study. In other words, starting from a diagnosis of a particular situation, one should derive actions of management to resolve the problems discovered. The step from diagnosis to therapy may not be as straightforward as one might wish. Clearly, it would be beneficial to partition the variables in the business model into those that can be influenced by management and those that cannot. If an unfavorable development is explained by a variable that cannot be controlled by management, then other ways will have to be found to resolve this problem. The integration of diagnosis with other elements of the managerial decision making process defines many new research problems, which will hopefully be addressed in the future.

Apart from the applicability of our diagnostic methods to the diagnosis of business performance, we also expect possibilities for applications in auditing. The auditing task can be described as providing a reasonable assurance that the financial statements of a company represent a fair picture of its operating results and financial position. To perform the auditing task, an accurate prediction of the company's financial statements is required. This prediction takes the role of "normal" behaviour in the diagnostic process. An important issue in auditing research is whether a discrepancy between predicted and actual behaviour is to be considered significant, or "material" in auditing terms. Explanation of such discrepancies could be performed by the explanation procedure for complete information diagnosis.

Appendix A: Case Study

In this appendix we describe a case study that we performed; its results have largely inspired the basic ideas underlying the theory of diagnosis and explanation of business performance developed in this thesis. The case study involves the analysis of the problem solving behaviour of a stock analyst at a large Dutch bank. The purpose was not to develop a system that would simulate human reasoning on this task in all respects, including its shortcomings and limitations. We *do* believe, however, that structuring a computer program's processes parallel to human decision-making processes, will make its results more acceptable to the user. The objective of this case study therefore was to take a real-life problem solving task, and to determine the essential elements of diagnostic reasoning in the business domain. This was done by analysing the think-aloud protocols of one analyst. Analysis of think-aloud protocols is nowadays an accepted technique in the knowledge acquisition phase of the development of knowledge-based systems. The technique originates from research in cognitive psychology. The function of think-aloud protocols, however, is different in *knowledge engineering* (the activity of designing and developing knowledge-based systems). In cognitive psychology research, the emphasis is on the faithful reproduction of the human problem solving process, including its obvious limitations. In knowledge engineering, the objective is to build a "well performing" computer model, which may result in an idealised representation of the problem-solving process. To illustrate this difference, we consider once more the research of Bouwman ([Bou78]), which was discussed in section 1.4.3. One of his findings was that in the identification of significant differences between norm and actual values, subjects were more likely to identify as significant differences that cross some multiple of ten. For example, a subject would consider the difference between 54 and 61 significant, whereas the difference between 52 and 59 would not be considered significant. Since there is no rationale for such behaviour, it is not desirable from a *knowledge engineering* point of view to model it.

We point out that the information obtained from the think-aloud protocols is not detailed enough to lead to the theory we developed in chapters 3 and 4 in an unambiguous way. The "operationalisation" of the general ideas obtained from this case study, has been based primarily on theoretical considerations, including results from other researchers, rather than information from the protocols.

This appendix is organized as follows. In the next section we give a short description of the task of a stock analyst at a large Dutch bank, the ABN-AMRO bank. Thereafter, we present an analysis of the protocols that resulted from a think-aloud experiment that we conducted. Finally, we draw a number of conclusions from this case study.

Task description

It has already been remarked in section 3.2 that the diagnosis and assessment of business performance can take place from different angles, depending on the viewpoint of the analyst. For example, a manager of a company, and a shareholder have different interests, and therefore will emphasize different aspects of the company's performance.

For the stock analyst, who evidently evaluates a company from the viewpoint of a shareholder, the ultimate assessment of a firm is the *expected return* on its shares. Expected return is calculated using the so-called Dividend Discount Model (D.D.M.),

$$P_0 = \sum_{t=1}^{\infty} \frac{D_t}{(1+k)^t},$$

where P_0 denotes the current price of a share, D_t denotes the expected dividend in period t, and k denotes the expected return. By making a forecast of future dividends D_t, k is implicitly estimated by the stock analyst. It is evidently impossible to make explicit forecasts of *all* future dividend payments, and therefore models are applied that assume, for example, a constant growth rate of the dividend payments. We will not go into the details of different models for forecasting expected return of a share. It suffices to notice that stock analysts have to make predictions of future dividend payments of "their" companies. Since dividend payments are usually determined by profits, this involves the prediction of future profits.

A stock analyst also has to make an assessment of the "quality" of a company. The quality of a company is indicated by the letters A, B, C, or D, and provides a qualitative indication of the risk that a particular stock will have a sharp price fall. On this scale, an A denotes the highest assessment, or alternatively, the lowest risk. The risk assessment is made on the basis of what are called *fundamental* properties of a company, such as its financial position, size, cost structure, and quality of management. The expected return and quality rating of a company are supplied by the analyst to another department of the bank, which is responsible for the composition of optimal investment portfolios.

Apart from predicting future profits and making a quality assessment, the task of an analyst is to write reports concerning the position of a company. Such reports are communicated to internal and external clients of the bank. The motives for writing a report are manifold, e.g. the issue of new stock, the appearance of an annual report, or the occurrence of developments that cause a substantial change in the position of a company. Depending on the type of the report, it may contain information such as description of the firm's activities, historical development of the firm, description of activities per sector, analysis of the financial position, and expectations concerning future market developments.

Sources of information that are used by the analyst to perform these tasks are: company annual reports, industry information, newspapers, magazines, electronic data-bases (such as Datastream and First-call), and personal contacts with the company.

As is often the case with real-world problem solving processes, the problem solving process of a stock analyst consists of a number of "generic" processes. Even a shallow analysis of the task concerned suggests that it involves three generic problem solving processes: prediction, assessment, and diagnosis. It is not our purpose to build a conceptual model of the problem solving process of a stock analyst. Rather, we are interested in determining the essential knowledge elements and reasoning steps with respect to the diagnosis of company performance.

Analysis of the think-aloud protocol

In order to analyse the problem solving process of the stock analyst, a number of information sources have been used. Firstly, the bank's literature on stock analysis was studied in order to get a theoretical basis for interviewing stock analysts. The main source of information was a think-aloud protocol, resulting from a case study solved by one analyst, henceforth called A1.

The daily task of A1 was to analyse Dutch construction companies that had a quotation on the stock exchange. In order to analyse A1's reasoning processes, a case study involving five German construction companies was developed. The reason to choose German instead of Dutch companies was to get better insight into the basic principles on which reasoning is based. Since the subject already had substantial knowledge about Dutch construction companies, too many conclusions would already be known beforehand. This would clearly limit the reasoning process. Since the German construction companies were not known beforehand to A1, this would force him to go "back to the basics".

The case study concerned the following five construction companies: Dyckerhoff, Strabag, Philip Holzmann, Bilfinger & Berger, and Hochtief. It contained the annual reports of all companies for the years 1987, 1988, and 1989. Furthermore, the financial company data, i.e. balance sheets and income statements, had been entered into a computer spreadsheet. This spreadsheet file was used by A1 during the solution of the case, not only for consultation but also for computing new figures when needed. The complete file was saved to disk after the case study, so that the new figures computed during the problem solving process could be included in the analysis. The use of a spreadsheet program to perform the analysis was common practice for A1, so the conditions of the case study resembled normal circumstances in this respect. A1 was asked to make an evaluation of the performance of the five companies involved, and to "think aloud" during this process. The think-aloud protocols were tape recorded and transcribed in full. The transcriptions were then analysed in order to discover the knowledge elements and reasoning used by the analyst for the diagnosis of a company's performance. The only preconceptions we had about diagnostic reasoning, when we started the analysis, was that it should involve the explanation of abnormal behaviour.

In the following a number of excerpts from the case transcription are presented, that provide some clear instances of applications of norms and explanation of differences between norm value and actual value. Excerpt 1 was recorded while A1 was scanning the financial data of Bilfinger & Berger in the worksheet.

Bilfinger & Berger	1988	1989
short term debt (% of balance sheet total)	38.4	45.3
short term debt (absolute)	745667	1067907

Table A.1: Data for excerpt 1

Excerpt 1:

Short term debt is going fast, let's have a look why that is. Yes, this is fairly important, short debt is about 45% of the balance sheet total in '89, it's increasing fairly rapidly, in percentage, absolute percentage. Have a look what the share is...yes this is going fast. This is another issue, have a look at it later.

This excerpt shows a clear example of problem identification. The analyst notices that short term debt of Bilfinger & Berger is increasing rapidly and that it amounts to 45% of the balance sheet total in 1989 (see table A.1). This indicates the application of a historical norm, i.e. the 1989 value is compared to the 1988 value, and a norm concerning the percentage of balance sheet total, although it is not stated what would be considered a normal percentage. Excerpt 2 shows a more elaborate example of the application of norms.

Excerpt 2:

Good current, good quick, actually very good, both of them. Here you can *see, this is strange,... current ratio equals current assets divided by current liabilities, well basically the higher the better. Quick ratio is the same except that you don't include the inventories in the current assets. You can see the difference, they're both very high; current ratio is about 3, the difference is very small. This means that the company hardly has any inventories. That is very strange for a construction company. This suggest they use a different bookkeeping method with respect to inventories because... When a construction company builds something, the project is usually divided into a number of terms, of phases, and depending on the arrangements with the commissioner, this is booked as inventory. A construction company builds something and it's 31 december and it's only half finished. Legally it's still his property. He may have received some payments from the commissioner, but legally it's still the property of the construction company, so this appears on the balance sheet as inventory. The amount that has already been paid is subtracted from this amount. If the agreement is: the commissioner pays every three months, and the last payment was on the* 1^{st} *of november, and today it's the* 31^{st} *of december, this means you have built for two months that have not been paid; this amount will be booked as inventory. This means that a construction company usually has a fairly large amount of inventories, and in this case we observe that there is hardly any inventory. In other words, how can this be explained?*

Hochtief	1988	1989
current ratio	2.91	2.80
quick ratio	2.75	2.56

Table A.2: Data for excerpt 2

This excerpt contains a rather elaborate justification of why a particular norm for the variable *inventory* applies. From the small difference between current ratio and quick ratio, shown in table A.2, the analyst *infers* that Hochtief hardly has any inventories. Since it is normal for a construction company to have a rather large amount of inventories (the norm value is not specified explicitly by A1), this discrepancy calls for an explanation. Notice that A1 evidently has "background" knowledge that enables him to give an elaborate justification of the norm for inventories of construction companies. Such knowledge can not be represented in the norm model or business model as described in sections 3.2 and 3.3 respectively.

In excerpt 3, we show a clear example of the way the analyst looks for explanations for abnormal behaviour. It was recorded after A1 had computed the solvency ratio in 1988 and 1989, for all five companies, and put them next to each other in the worksheet.

Excerpt 3:

Secondly, solvency is also important to profitability, because a very low solvency ratio inhibits your access to loan capital. If you have a low proportion of equity capital, then suppliers of loan capital will find you risky. That's why they'll demand a higher interest rate. Due to the financial leverage effect, this will directly affect your return on equity. In general you can set a certain criterion, a Dutch construction company should have a solvency ratio of about 0.3. Germany, I don't know, looking at these figures I would say 0.3 is normal in Germany too. In that case Holzmann clearly deviates, the others are o.k. Bilfinger is a little bit low, I would say it like that, but not discomforting. In this sequence, Hochtief is clearly the best, Holzmann is too low, let's have a look why. Then you start looking at the composition...and then you have this problem of the high provisions with the Germans. Provisions are made for risks you know you will encounter in the future, but you don't know how high they will be. If you want to assess how risky a company is,...those provisions have not been claimed yet and that makes a difference in case of, ultimately, a bankruptcy. Then all creditors will be lined up, if your external capital, all 36.8, were direct debt to creditors, then there will be for 36.8 million worth of creditors lined up. With provisions this is not the case. This diminishes the risk a little bit, because you know that only the external capital of creditors will be taken away first. If there is still a large amount left, then you will receive some money as a shareholder. So *you would have to include this consideration. Yes you can see a very low equity capital, this low ratio of equity to debt could be due to a low equity capital or a high debt or both. Well in this case you can see that debt does not deviate, so it's a matter of low equity capital. Well that gives a negative assessment for two reasons.*

	Solvency	
Company	1988	1989
Strabag	0.25	0.28
Hochtief	0.47	0.49
Holzmann	0.12	0.17
Dyckerhoff	0.42	0.42
Bilfinger	0.23	0.26

Table A.3: Data for excerpt 3

In this excerpt, the norm value that is applied to the variable *solvency* is stated explicitly: since the analyst is not very familiar with German construction companies, he applies the norm of their Dutch counterparts. Looking at the companies considered in this case study, he concludes that Holzmann is significantly below the norm; see table A.3. Since solvency is defined as equity divided by debt, the analyst hypothesizes that a low solvency could be due to either a low equity or a high debt, or both. Since Holzmann's debt does not deviate, it is inferred that a low equity capital has caused the low solvency.

Conclusions

The case study described in this appendix gives rise to a number of conclusions.

Firstly, the analyst selects "events" that should be explained, by comparing a variable's value against some norm value. We saw several examples of this in the protocol excerpts. The norm that is used is either a historical value or some kind of mean value of a group of companies.

Secondly, not *every* deviation between norm value and actual value results in an attempt at explanation. We saw a clear example of this in excerpt 3, where actually *all* companies deviated from the norm value of 0.3 for solvency. Only the deviation of Holzmann induced an attempt at explanation by A1.

Thirdly, we conclude that the explanation of a variable's deviation from the norm takes place by looking at its "influencing" variables. For example, in excerpt 3 the analyst explained Holzmann's low solvency by looking at its defining variables, i.e. equity and debt. However, not every deviation from the norm is considered to be an explanatory cause. This also depends on the amount of deviation.

All these findings, which are largely in accordance with the results of Bouwman ([Bou78]), have been incorporated into our method for explanation and diagnosis. We regarded them as *minimum requirements* that should be met by our method. However, the results of the protocol analysis are in no sense detailed enough to lead unambiguously to an operational method of diagnosis. The "operationalisation" into a method of diagnosis, has been inspired by theoretical considerations.

The think-aloud protocol and a number of additional interviews have also served as the source for the development of a tool to support the analyst in the assessment of Dutch construction companies. In building this tool, the objective was not to imitate the human analyst but rather to support his decision making while maintaining a close analogy to his way of reasoning. The business model was, to a large extent, based on a spreadsheet model that was already used by the analyst to evaluate the financial data of construction companies. The tool has not been formally evaluated. Informal evaluation showed that the analysts considered the program to be a useful extension of the spreadsheet software they were already using. However, as was to be expected, the analysts also recognized that the program covers only a part of their tasks.

Bibliography

[AH86] Chidanand Apté and Se June Hong. Using qualitative reasoning to understand financial arithmetic. Technical Report RC 11946, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 1986.

[AMCP88] Nassar H. Ata Mohammed, James F. Courtney, and David B. Paradice. A prototype DSS for structuring and diagnosing managerial problems. *IEEE Transactions on systems, man, and cybernetics*, 18(6):899–907, 1988.

[Be84] Daniel G. Bobrow (ed.). Special issue on qualitative reasoning. *Artificial Intelligence*, 24, 1984.

[Ber92] Ron J. Berndsen. *Qualitative Reasoning and Knowledge Representation in Economic Models.* PhD thesis, Tilburg University, 1992.

[Bon72] John W. Bonge. Problem recognition and diagnosis: basic inputs to business policy. *Journal of business policy*, pages 45–53, 1972.

[Bou78] Marinus J. Bouwman. *Financial diagnosis: a cognitive model of the processes involved.* PhD thesis, Carnegie-Mellon University, 1978.

[Bou83] Marinus J. Bouwman. Human diagnostic reasoning by computer: an illustration from financial analysis. *Management Science*, 29(6):653–672, 1983.

[Bou84] M. Boutillier. Reading macroeconomic models and building causal structures. In J.P. Ancot, editor, *Analysing the structure of econometric models*, pages 35–48. Martinus Nijhoff Publishers, The Hague, 1984.

[BS84] B.G. Buchanan and E.H. Shortliffe. *RULE-BASED EXPERT SYSTEMS, The MYCIN experiments of the Stanford Heuristic Programming Project.* Addison Wesley Publishing Company, Reading, MA, 1984.

[CKRP73] A. Colmerauer, H. Kanoui, P. Roussel, and R. Pasero. Un systeme de communication homme-machine en francais. Technical report, Groupe de Recherche en Intelligence Artificielle, Université d'Aix-Marseille, 1973.

[Col90] Alain Colmerauer. An introduction to Prolog-III. *Communications of the ACM*, 33(7):69–90, 1990.

[CPAM87] James F. Courtney, David B. Paradice, and Nassar H. Ata Mohammed. A knowledge-based DSS for managerial problem diagnosis. *Decision Sciences*, 18:373–399, 1987.

[CS83] B. Chandrasekaran and Mittal Sanjay. Deep versus compiled knowledge approaches to diagnostic problem-solving. *International Journal of Man-Machine Studies*, 19:425–436, 1983.

[CT90] Luca Console and Pietro Torasso. Hypothetical reasoning in causal models. *International journal of intelligent systems*, 5:83–124, 1990.

[DF90] H. Daniels and A. Feelders. Model-based diagnosis of business performance. In *Tenth international workshop on expert systems and their applications*, pages 155–164, Avignon, 1990.

[DF91] H.A.M. Daniels and A.J. Feelders. Combining qualitative and quantitative methods for model-based diagnosis of firms. In Madan G. Singh and Louise Travé-Massuyès, editors, *Decision Support Systems and Qualitative Reasoning*, pages 255–260. North-Holland, Amsterdam, 1991.

[DF92] H.A.M. Daniels and A.J. Feelders. Explanation and diagnosis in business assessment. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(2):397–402, 1992.

[DH88] Randall Davis and Walter Hamscher. Model-based reasoning: troubleshooting. In H. E. Shrobe, editor, *Exploring Artificial Intelligence: survey talks from the national conferences on AI*, pages 297–346. Morgan Kaufmann, 1988.

[DHS+88] Mehmet Dincbas, P. van Hentenryck, H. Simonis, A. Aggoun, T. Graf, and F. Berthier. The constraint logic programming language CHIP. In *Proceedings of the international conference on fifth generation computer systems (FGCS '88)*, Tokyo, Japan, 1988.

[DKB84] Johan De Kleer and John Seely Brown. A qualitative physics based on confluences. *Artificial Intelligence*, 24:7–83, 1984.

[dVR78] Pieter F. de Vries Robbé. *Medische besluitvorming, een aanzet tot formele geneeskunde.* PhD thesis, Groningen University, 1978.

[EH82] H.J. Einhorn and R.M. Hogarth. Prediction, diagnosis and causal thinking in forecasting. *Journal of Forecasting*, 1(1):23–36, 1982.

[EN68] C. William Emory and Powell Niland. *Making Management Decisions.* Houghton Mifflin Company, New York, 1968.

[Fee92] Ad Feelders. A model of explanation for financial knowledge-based systems. *Computer Science in Economics and Management*, 5:119–132, 1992.

[FL90] A. Farley and K. Lin. Qualitative reasoning in economics. *Journal of Economic Dynamics and Control*, 14, 1990.

[For84] Kenneth D. Forbus. Qualitative process theory. *Artificial Intelligence*, 24:85–168, 1984.

[For87] Kenneth J. Fordyce. Looking at worksheet modeling through expert system eyes. In G. Silverman, editor, *Expert systems for business*, pages 246–285. Addison-Wesley, 1987.

[Gen84] Michael R. Genesereth. The use of design descriptions in automated diagnosis. *Artificial Intelligence*, 24:411–436, 1984.

[Gra80] C.W.J. Granger. Testing for causality, a personal viewpoint. *Journal of Economic Dynamics and Control*, 2:329–352, 1980.

[Ham90] Walter Hamscher. Explaining unexpected financial results. Paper presented at the AAAI spring symposium on abduction, 1990.

[HBD86] Peter E. Hart, Amos Barzilay, and Richard O. Duda. Qualitative reasoning for financial assessments: a prospectus. *The AI Magazine*, 7:62–68, 1986.

[Hel78] Erich A. Helfert. *Techniques of Financial Analysis.* Prentice-Hall, 1978.

[Hem65] Carl G. Hempel. Aspects of scientific explanation. In *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, pages 331–496. The Free Press, New York, 1965.

[Hes84] Germund Hesslow. Explaining differences and weighting causes. *Theoria*, 49:87–111, 1984.

[Hum89] Paul W. Humphreys. *The Chances of Explanation.* Princeton University Press, Princeton, New Jersey, 1989.

[Hux88] Avon Huxor. Explanation and understanding: an architecture for economic elaboration. Technical Report TR-LP-26, ECRC, München, Germany, 1988.

[ITH80] H. Ingham and L. Taylor Harrington. *Interfirm comparison.* Heinemann, London, 1980.

[Iwa89] Yumi Iwasaki. Qualitative physics. In *The Handbook of Artificial Intelligence 4*. Addison-Wesley, Reading, 1989.

[Jac89] Peter Jackson. Applications of nonmonotonic logic to diagnosis. *The Knowledge Engineering Review*, 4(2):97–117, 1989.

[JMSY92] Joxan Jaffar, Spiro Michaylov, Peter J. Stuckley, and Roland H.C. Yap. The CLP(R) language and system. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 14(3):339–395, 1992.

[JPW81] Phillip Judd, Charles Paddock, and James Wetherbe. Decision impelling differences: an investigation of management by exception reporting. *Information & Management*, 4:259–267, 1981.

[KK84] Benjamin Kuipers and Jerome P. Kassirer. Causal reasoning in medicine: analysis of a protocol. *Cognitive Science*, 8:363–385, 1984.

[Kos89] Donald W. Kosy. Applications of explanation in financial modeling. In L.E. Widman, Kenneth A. Loparo, and Norman R. Nielsen, editors, *Artificial Intelligence, Simulation and Modeling*, pages 487–509. John Wiley & Sons, 1989.

[Kow74] Robert A. Kowalski. Predicate logic as a programming language. In *IFIP 74*, pages 569–574, 1974.

[Kui86] Benjamin Kuipers. Qualitative simulation. *Artificial Intelligence*, 29:289–338, 1986.

[KW84] Donald W. Kosy and Ben P. Wise. Self-explanatory financial planning models. In *Proceedings of AAAI-84*, pages 176–181, Los Altos, CA, 1984. Morgan Kaufmann.

[Llo84] J.W. Lloyd. *Foundations of logic programming.* Springer-Verlag, Berlin, 1984.

[MRT76] Henry Mintzberg, Duru Raisinghani, and André Théorêt. The structure of "unstructured" decision processes. *Administrative Science Quarterly*, 21:246–275, 1976.

[Pou69] William F. Pounds. The process of problem finding. *Industrial Management Review*, 11(1):1–19, 1969.

[PSS84] Ramesh S. Patil, Peter Szolovits, and William B. Schwartz. Causal understanding of patient illnes in medical diagnosis. In William J. Clancey and Edward H. Shortliffe, editors, *Readings in Medical Artificial Intelligence*, pages 339–360. Addison-Wesley, 1984.

[Rau88] Norbert Rauh. *Wissensbasierte Systeme zur Unternehmensdiagnose auf der Grundlage von Jahresabschlussdaten und Branchenvergleichswerten in der Steuerkanzlei.* PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 1988.

[Rei87] Raymond Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–96, 1987.

[RNW84] James A. Reggia, Dana S. Nau, and Pearl Y. Wang. Diagnostic expert systems based on a set covering model. In M.J. Coombs, editor, *Developments in expert systems*, pages 35–58. Academic Press, 1984.

[Sim86] Reid Simmons. Commonsense arithmetic reasoning. In *Proceedings of AAAI-86*, pages 118–124, Los Altos, CA, 1986. Morgan Kaufmann.

[Ste88] Luc Steels. Components of expertise. Technical Report AI MEMO 88-16, VUB AI-LAB, Brussels, July 1988.

[vF80] Bas van Fraassen. The pragmatics of explanation. In *The Scientific Image*, pages 97–157. Clarendon Press, Oxford, 1980.

[VH89] Pascal Van Hentenryck. *Constraint satisfaction in logic programming.* The MIT Press, Cambridge, MA, 1989.

[VH91] Pascal Van Hentenryck. Constraint logic programming. *The Knowledge Engineering Review*, 6(3):151–194, 1991.

[WC89] Irene M.Y. Woon and Peter Coxhead. Qualitative modelling in financial analysis. In Louis F. Pau, J. Motiwalla, Y.H. Pao, and H.H. Teh, editors, *Artificial intelligence in economics, banking and management*, pages 77–86. Elsevier Science Publishers (North-Holland), 1989.

[Whi65] Morton White. *Foundations of Historical Knowledge*. Harper & Row, New York, 1965.

[Wil84] Brian C. Williams. Qualitative analysis of mos circuits. *Artificial Intelligence*, 24:281–346, 1984.

[WKAS78] S.M. Weiss, C.A. Kulikowski, S. Amarel, and A. Safir. A model-based method for computer-aided medical decision making. *Artificial Intelligence*, 11, 1978.

Samenvatting

De formalisering van diagnostisch redeneren is een deelgebied van de artificiële intelligentie (AI) dat de laatste jaren intensief bestudeerd is. Hierbij definiëren we diagnose als het vinden van de beste verklaring voor waargenomen abnormaal gedrag van een systeem. Met name model-gebaseerde, als tegenhanger van "heuristische classificatie", benaderingen zijn uitgebreid onderzocht, vooral vanwege hun verondersteide superioriteit in het oplossen van problemen en het geven van verklaringen.

Het grootste deel van het onderzoek op het gebied van diagnose heeft zich, zo niet expliciet dan toch impliciet, gericht op medische diagnose en diagnose van technische apparaten. Dit heeft zijn gevolgen gehad voor de kennisrepresentaties en de daarmee samenhangende redeneermethoden die hiertoe zijn ontwikkeld.

Bij de diagnose van electronische schakelingen bijvoorbeeld, wordt het systeem meestal weergegeven door middel van een verzameling eerste-orde logica uitspraken die een beschrijving geeft van het gedrag van de componenten, en de manier waarop ze met elkaar zijn verbonden. In het medische domein treft men veelal causale modellen aan, die een oorzaak-gevolg relatie beschrijven tussen "ziekte toestanden". In overeenstemming met de kennis die in dit domein voorhanden is, zijn deze causale modellen van kwalitatieve aard.

In deze studie hebben we ons daarentegen toegelegd op de formalisering van diagnostisch redeneren in het bedrijfseconomische domein, in het bijzonder de diagnose van ondernemingsresultaten. Hierbij hebben we het begrip "verklaring" als centraal uitgangspunt genomen. In hoofdstuk 2 ontwikkelen we een causale theorie van verklaringen die in staat is om te gaan met de kwantitatieve verschijnselen die veelvuldig voorkomen in het bedrijfseconomische domein.

Met betrekking tot de kennisrepresentatie kan onze formalisering worden gekarakteriseerd als model-gebaseerd. Afgezien van de voordelen die aan model-gebaseerd redeneren worden toegeschreven, is het ook de meest natuurlijke kennisrepresentatie in dit domein. Dit komt doordat een groot gedeelte van de benodigde kennis betrekking heeft op de relatie tussen financiële en operationele variabelen. Deze kennis is vaak al aanwezig in de vorm van een kwantitatief model dat gebruikt wordt om waarden te *berekenen* in plaats van te *verklaren*. Daarentegen is kennis van "echte" oorzaak gevolg relaties vaak van kwalitatieve aard.

In hoofdstuk 4 ontwikkelen we twee benaderingen van diagnose, één gebaseerd op de veronderstelling dat volledige informatie voorhanden is, en één gebaseerd op onvolledige informatie. In het tweede geval zijn we gedwongen tot kwalitatief redeneren teneinde een eindig aantal verklarende hypothesen te krijgen. Onvolledige informatie lijkt een noodzakelijke eigenschap van diagnoseproblemen in andere domeinen. Naar onze mening is dit niet het geval in het bedrijfseconomische domein, waar de selectie van significante invloeden in het geval van volledige informatie ook als diagnose wordt beschouwd. In de praktijk zouden beide benaderingen in één systeem kunnen worden gecombineerd, waarbij kwantitatief en kwalitatief redeneren worden afgewisseld al naargelang de beschikbare informatie. Vergelijking van de verklaringen die onze methode oplevert met verklaringen uit een tekstboek, laat zien dat ze in grote mate met elkaar overeenkomen.

In hoofdstuk 5 bespreken we de implementatie van twee diagnostische programma's in respectievelijk de logische programmeertalen PROLOG en CHIP. Deze programmeertalen bleken de nodige flexibiliteit te bezitten om het ondernemingsmodel enerzijds te beschouwen als declaratieve kennis waarmee symbolische manipulaties uitgevoerd kunnen worden, en anderzijds te gebruiken als een rekenmodel.

De diagnostische methoden die we hebben ontwikkeld zijn enerzijds gebaseerd op het - weinige - werk dat is verricht op het gebied van automatische diagnose in het bedrijfseconomische domein, en anderzijds op een case study die we hebben uitgevoerd. In de appendix beschrijven we deze case study die betrekking heeft op de analysetaak van een aandelenanalist bij de AMRO-bank. We hebben een analyse gemaakt van de hardopdenk protocollen van deze aandelenanalist tijdens het evalueren van de resultaten van een aantal Duitse bouwondernemingen. Deze analyse bracht een aantal eigenschappen van probleemidentificatie en het geven van verklaringen door de financieel analist aan het licht. Deze eigenschappen zijn verwerkt in het formele model van diagnose, zowel waar het het toepassen van normen als waar het de verklaring van afwijkingen van die normen betreft.

Dankwoord

Op deze plaats wil ik een aantal personen bedanken die op de een of andere wijze hebben bijgedragen aan het tot stand komen van dit proefschrift.

Op de eerste plaats ben ik dank verschuldigd aan mijn dagelijkse begeleider en copromotor Hennie Daniels, voor zijn intensieve begeleiding en betrokkenheid bij het onderzoek.

Ik dank mijn promotor Prof. J.P.C. Kleijnen, die dit onderzoek altijd krachtig heeft gesteund, en Pieter Spierings, destijds aandelenanalist bij de AMRO-bank, voor zijn enthousiaste medewerking aan dit onderzoek.

Tevens ben ik dank verschuldigd aan de mensen van de *Logic Programming Group* van het ECRC, in het bijzonder Alexander Herold, Mehmet Dincbas, Françoise Berthier en Helmut Simonis. Zij hebben ertoe bijgedragen dat mijn verblijf als gastonderzoeker zowel leerzaam als plezierig was.

Tenslotte dank ik mijn ex-collega's en kamergenoten Ron Berndsen, Johan Broek en William Verkooijen. Ze weten wel waarvoor.